

1N-61-CR

36982

P-38

# MULTI-STAGE DECODING FOR MULTI-LEVEL BLOCK MODULATION CODES

Technical Report

to

NASA

Goddard Space Flight Center

Greenbelt, Maryland 20771

Grant Number NAG 5-931

Report Number NASA 91-002

Shu Lin

Principal Investigator

Department of Electrical Engineering

University of Hawaii at Manoa

Honolulu, Hawaii 96822

August 27, 1991

(NASA-CR-188759) MULTI-STAGE DECODING FOR  
MULTI-LEVEL BLOCK MODULATION CODES (Hawaii  
Univ.) 38 p CSCL 09B

N91-30747

Unclas  
G3/61 0036982

# MULTI-STAGE DECODING FOR MULTI-LEVEL BLOCK MODULATION CODES

SHU LIN

TADAO KASAMI

## Abstract

In this paper, we investigate various types of multi-stage decoding for multi-level block modulation codes, in which the decoding of a component code at each stage can be either soft-decision or hard-decision, maximum likelihood or bounded-distance. Error performance of codes is analyzed for a memoryless additive channel based on various types of multi-stage decoding, and upper bounds on the probability of an incorrect decoding are derived. Based on our study and computation results, we find that, if component codes of a multi-level modulation code and types of decoding at various stages are chosen properly, high spectral efficiency and large coding gain can be achieved with reduced decoding complexity. In particular, we find that the difference in performance between the suboptimum multi-stage soft-decision maximum likelihood decoding of a modulation code and the single-stage optimum decoding of the overall code is very small, only a fraction of dB loss in SNR at the probability of an incorrect decoding for a block of  $10^{-6}$ . Multi-stage decoding of multi-level modulation codes really offers a way to achieve the best of three worlds, bandwidth efficiency, coding gain and decoding complexity.

## 1. Introduction

Multi-level method is a powerful technique for constructing bandwidth efficient modulation codes [1-7]. This method allows us to construct modulation codes systematically to achieve high spectral efficiency and large coding gain from component codes (binary or nonbinary, trellis or block) in conjunction with proper bits-to-signal mapping through signal set partitioning. If the component codes are chosen properly, the resultant multi-level code not only has good minimum squared Euclidean distance but is also rich in structural properties such as: linear (or regular) structure, phase symmetry, and trellis structure [4, 7, 8].

A major advantage of multi-level modulation codes is that these codes can be decoded in multiple stages with component codes decoded sequentially stage by stage and with decoded information passed from one stage to the next stage. Since the component codes are decoded one at a time, it is possible to take advantage of the structure of each component code to simplify the decoding complexity and reduce the number of computations at each decoding stage. As a result, the overall complexity and the number of computations needed for decoding a multi-level modulation code will be greatly reduced. Decoding of a component code at each stage can be either soft-decision or hard-decision, maximum likelihood or bounded-distance. If component codes and the types of decoding at various stages are chosen properly, high spectral efficiency and large coding gain (or high reliability) can be achieved with reduced decoding complexity. Multi-stage decoding for multi-level modulation codes really offers a way to achieve the best of three worlds, bandwidth efficiency, coding gain and decoding complexity.

In this paper, we investigate and analyze various types of multi-stage decoding for multi-level block modulation codes, particularly the multi-stage soft-decision maximum likelihood decoding, multi-stage hard-decision maximum likelihood decoding, multi-stage bounded-distance decoding and hybrid multi-stage decoding. The organization of the paper is as follows. In Section 2, we provide a general formulation for multi-level block modulation codes in terms of component codes over substrings of labeling symbols. In Section 3, soft-decision maximum likelihood, hard-decision maximum likelihood and bounded-distance decodings for a component code of a multi-level modulation code are devised. In Section 4, the error performance of multi-level block modulation codes for a memoryless additive channel is analyzed based on various types of multi-stage decoding, and upper bounds on the probability of an

incorrect decoding are derived. Finally, error performance and coding gain of some specific bandwidth efficient multi-level modulation codes are computed and simulated for various types of multi-stage decoding. From our computation and simulation results, we find that multi-stage decoding for multi-level modulation codes provides an excellent trade-off between error performance (or coding gain) and decoding complexity.

## 2. Multi-Level Block Modulation Codes

Construction of multi-level modulation codes consists of six basic steps: (1) selection of a signal set  $S$ ; (2) labeling of signal points by strings of labeling symbols through signal set partitioning; (3) segmentation of signal labels into sub-labels; (4) selection (or construction) of component codes over the sub-labels; (5) combining component codes by concatenating the sub-labels to form a multi-level codes; and (6) label-to-signal mapping to form a multi-level modulation code.

Consider a signal set (or constellation)  $S$  with  $2^\ell$  signal points where  $\ell$  is a positive integer. Suppose the signal points in  $S$  are labeled by binary strings of length  $\ell$  through a proper set partitioning process [7,9–11]. Then the label set  $L$  for  $S$  is of the form:

$$L \triangleq \{a_1 a_2 \cdots a_\ell : a_i \in \{0, 1\} \text{ for } 1 \leq i \leq \ell\}.$$

Let  $\sigma(\cdot)$  be the mapping defined on  $L$  such that  $\sigma(a_1 a_2 \cdots a_\ell)$  gives a unique signal point  $s$  in  $S$ . The labeling  $(L, \sigma)$  (or simply  $L$ ) is said to have  $\ell$  levels or length  $\ell$ .

For  $v$  and  $v'$  in  $L$ , let  $d(v, v')$  denote a distance measure between two signal points,  $s$  and  $s'$  in  $S$ , labeled  $v$  and  $v'$  respectively. The distance measure is assumed to have the property that  $d(v, v') = d(v', v)$  and  $d(v, v') = 0$  if and only if  $v = v'$ . For a positive integer  $n$ , let  $X^n$  denote the set of all  $n$ -tuples over a set  $X$ . We extend the domain of  $d$  as follows: For two  $n$ -tuples,  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  and  $\mathbf{v}' = (v'_1, v'_2, \dots, v'_n)$  over  $L$ , define

$$d(\mathbf{v}, \mathbf{v}') \triangleq \sum_{j=1}^n d(v_j, v'_j). \quad (2.1)$$

For a nonempty subset  $C$  of  $L^n$ , define the minimum distance of  $C$  with respect to measure  $d$ , denoted  $D[d, C]$ , as follows:

$$D[d, C] \triangleq \min \{ d(\mathbf{v}, \mathbf{v}') : \mathbf{v}, \mathbf{v}' \in C \text{ and } \mathbf{v} \neq \mathbf{v}' \}. \quad (2.2)$$

(If  $|C| = 1$ , then  $D[d, C]$  is defined as infinity.) For simplicity, we use  $D[C]$  to denote  $D[d, C]$  whenever there is no ambiguity.

Suppose a labeling  $L$  of  $\ell$  levels for a signal set  $S$  and a distance measure  $d$  are given. For constructing a general multi-level code over  $L$ , we must segment the labeling into sub-labelings and choose the starting symbol position of each sub-labeling. Let  $m$  be a positive integer not greater than  $\ell$ , and let  $j_1, j_2, \dots, j_{m+1}$  be  $m+1$  integers such that

$$1 = j_1 < j_2 < \dots < j_m < j_{m+1} = \ell + 1. \quad (2.3)$$

For  $1 \leq i \leq m$ , let  $\ell^{(i)}$  be defined as

$$\ell^{(i)} \triangleq j_{i+1} - j_i,$$

and let  $L^{(i)}$ , called the  $i$ -th sub-labeling, denote the set of substrings from the  $j_i$ -th symbol to the  $(j_{i+1} - 1)$ -th symbol of strings in  $L$ , i.e.,

$$L^{(i)} \triangleq \{ a_{j_i} a_{j_i+1} \dots a_{j_{i+1}-1} : a_h \in \{0, 1\} \text{ for } j_i \leq h < j_{i+1} \}. \quad (2.4)$$

Concatenating  $L^{(1)}$  to  $L^{(m)}$ , we obtain

$$L = L^{(1)} L^{(2)} \dots L^{(m)}.$$

Consider an  $n$ -tuple  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  over  $L$ . For  $1 \leq j \leq n$ , the  $j$ -th component  $v_j$  of  $\mathbf{v}$  can be expressed as the following concatenation of substrings in  $L^{(1)}$  to  $L^{(m)}$ :

$$v_j = v_j^{(1)} v_j^{(2)} \dots v_j^{(m)}$$

where  $v_j^{(i)} \in L^{(i)}$  for  $1 \leq i \leq m$ . For  $1 \leq i \leq m$ , we form the following  $n$ -tuple over  $L^{(i)}$ :

$$\mathbf{v}^{(i)} = (v_1^{(i)}, v_2^{(i)}, \dots, v_n^{(i)}). \quad (2.5)$$

This  $n$ -tuple  $\mathbf{v}^{(i)}$  is called the  $i$ -th component  $n$ -tuple of  $\mathbf{v}$ , and  $\mathbf{v}$  is denoted as follows:

$$\mathbf{v} = \mathbf{v}^{(1)} \mathbf{v}^{(2)} \dots \mathbf{v}^{(m)}. \quad (2.6)$$

For  $1 \leq i \leq m$ , let  $C_i$  be a block code of length  $n$  over  $L^{(i)}$ . From  $C_1, C_2, \dots, C_m$ , we form a block code of length  $n$  over  $L$  as follows:

$$\begin{aligned} C &\triangleq C_1 C_2 \dots C_m \\ &= \{ \mathbf{v}^{(1)} \mathbf{v}^{(2)} \dots \mathbf{v}^{(m)} : \mathbf{v}^{(i)} \in C_i \text{ for } 1 \leq i \leq m \}. \end{aligned} \quad (2.7)$$

Such a code  $C$  is called an  $\ell$ -level code with  $m$  components, and  $C_i$  is called the  $i$ -th component code of  $C$ . If each component of a codeword in  $C$  is mapped into the corresponding signal point in the signal constellation  $S$ , we obtain a multi-level block modulation code.

For a distance measure  $d$  on  $L$  and  $1 \leq i \leq m$ , let  $d^{(i)}(v^{(i)}, v'^{(i)})$  with  $v^{(i)}$  and  $v'^{(i)}$  in  $L^{(i)}$  be defined as follows:

$$d^{(i)}(v^{(i)}, v'^{(i)}) \triangleq \min \left\{ d(v^{(1)} \dots v^{(i-1)} v^{(i)} v^{(i+1)} \dots v^{(m)}, v^{(1)} \dots v^{(i-1)} v'^{(i)} v^{(i+1)} \dots v^{(m)}) : \right. \\ \left. v^{(j)} \in L^{(j)} \text{ with } j = 1, \dots, i-1, i+1, \dots, m \text{ and } \right. \\ \left. v'^{(j)} \in L^{(j)} \text{ with } j = i+1, \dots, m \right\}. \quad (2.8)$$

For any distance measure  $d$ , a lower bound on the minimum distance  $D[d, C]$  of a multi-level code  $C$  is given by (2.9) [7], which unifies the previous bounds [1, 2, 4, 6, 12],

$$D[d, C] \geq \min_{1 \leq i \leq m} D[d^{(i)}, C_i]. \quad (2.9)$$

The equality in (2.9) holds if the following conditions (2.10) and (2.11) are satisfied:

- (1) For any  $v^{(j)}$  and  $v'^{(j)}$  in  $L^{(j)}$  with  $1 \leq j \leq m$  and any positive integer  $i$  not greater than  $m$ ,

$$d(v^{(1)} \dots v^{(i-1)} v^{(i)} v^{(i+1)} \dots v^{(m)}, v^{(1)} \dots v^{(i-1)} v'^{(i)} v^{(i+1)} \dots v^{(m)}) \\ = d(v'^{(1)} \dots v'^{(i-1)} v^{(i)} v'^{(i+1)} \dots v'^{(m)}, v'^{(1)} \dots v'^{(i-1)} v'^{(i)} v'^{(i+1)} \dots v'^{(m)}) \quad (2.10)$$

and

(2)

$$d^{(i)}(v^{(i)}, v'^{(i)}) = \min \left\{ d(v^{(1)} \dots v^{(i-1)} v^{(i)} v^{(i+1)} \dots v^{(m)}, v^{(1)} \dots v^{(i-1)} v'^{(i)} v^{(i+1)} \dots v^{(m)}) : \right. \\ \left. v^{(j)} \in L^{(j)} \text{ with } j = 1, \dots, i-1, i+1, \dots, m \right\}. \quad (2.11)$$

From (2.9), we see that to maximize  $D[d, C]$ , we need to form a labeling  $L$  for the signal set  $S$  which maximizes  $D[d^{(i)}, C_i]$  for all  $i$ .

For any  $i$  such that  $\ell^{(i)} = 1$ , the  $i$ -th sub-labeling is  $L^{(i)} = \{0, 1\}$ . Then, for any two  $n$ -tuples  $\mathbf{u}$  and  $\mathbf{v}$  over  $L^{(i)} (= \{0, 1\})$ ,

$$d^{(i)}(\mathbf{u}, \mathbf{v}) = d^{(i)}(0, 1) d_H(\mathbf{u}, \mathbf{v}), \quad (2.12)$$

where  $d_H(\mathbf{u}, \mathbf{v})$  denotes the Hamming distance between  $\mathbf{u}$  and  $\mathbf{v}$ . Hence, for  $\ell^{(i)} = 1$ ,  $C_i$  is a binary code and

$$D[d^{(i)}, C_i] = d^{(i)}(0, 1)\delta_i, \quad (2.13)$$

where  $\delta_i$  denotes the minimum Hamming distance of  $C_i$ .

Consider the special case for which  $m = \ell$ . Then for  $1 \leq i \leq \ell$ ,  $\ell^{(i)} = 1$  and  $L^{(i)} = \{0, 1\}$ . Consequently, an  $\ell$ -level code  $C$  is formed from  $\ell$  binary component codes,  $C_1, C_2, \dots$ , and  $C_\ell$ . For  $1 \leq i \leq \ell$ , define the  $i$ -th distance parameter  $d_i$  of  $L$  as

$$d_i \triangleq d^{(i)}(0, 1). \quad (2.14)$$

Let  $\delta_i$  be the minimum Hamming distance of component code  $C_i$  for  $1 \leq i \leq \ell$ . Then it follows from (2.9), (2.13) and (2.14) that the minimum distance of an  $\ell$ -level modulation code  $C$  with  $\ell$  binary component codes satisfies the following lower bound:

$$D[d, C] \geq \min_{1 \leq i \leq \ell} \delta_i d_i.$$

Since  $d_i \leq d_{i+1}$  for  $0 \leq i < \ell$ , we need to choose the component codes such that  $\delta_i \geq \delta_{i+1}$ . Therefore,  $C_1$  is the most powerful component code in terms of Hamming distance and  $C_\ell$  is the least powerful component code. An  $\ell$ -level code with  $\ell$  components is called a basic multi-level code. Most of the known block modulation codes [2,3,5,6,13] are basic multi-level codes.

Let  $M$  denote the integer  $2^\ell$ . For an  $M$ -QASK or  $M$ -PSK signal set, the squared Euclidean distance is used as the distance measure. For an  $M$ -QASK signal set, a binary labeling  $L$  of length  $\ell$  is chosen in such a way [9–11] that for  $1 < i \leq \ell$ ,

$$d_i = 2d_{i-1}. \quad (2.15)$$

This labeling is denoted  $L_{M\text{-QASK}}$ . For a binary string  $a_1 a_2 \dots a_j$ , let  $I(a_1 a_2 \dots a_j)$  denote the integer  $\sum_{i=1}^j a_i 2^{i-1}$  (for the null string  $\lambda$ ,  $I(\lambda) \triangleq 0$ ). For an  $M$ -PSK signal set with unit energy, the signal point in a 2-dimensional space labeled by a binary string  $u$  of length  $\ell$  is given by  $(\cos(2\pi I(u)/M), \sin(2\pi I(u)/M))$ , denoted  $s(u)$ , and the distance measure  $d$  between two binary strings  $u$  and  $v$  is given by the squared Euclidean distance between  $s(u)$  and  $s(v)$ , that is,

$$d(u, v) \triangleq 4 \sin^2(M^{-1} \pi (I(u) - I(v))). \quad (2.16)$$

This labeling is denoted  $L_{M\text{-PSK}}$ . It can be easily shown that if either  $i = m$  or  $\ell^{(i)} = 1$ , then for  $u$  and  $v$  in  $L_{M\text{-PSK}}^{(i)}$ ,

$$d^{(i)}(u, v) = 4 \sin^2(2^{j_i-1-\ell} \pi (I(u) - I(v))). \quad (2.17)$$

Hence the  $i$ -th distance parameter  $d_i$  of  $L_{M\text{-PSK}}$  is given by [9],

$$d_i = 4 \sin^2(2^{i-1-\ell} \pi). \quad (2.18)$$

Figure 1 shows an 8-PSK signal set with unit energy. Every signal point is labeled with a string of three bits,  $a_1 a_2 a_3$ . The distance parameters of  $L_{8\text{-PSK}}$  are:  $d_1 = 0.586$ ,  $d_2 = 2$ ,  $d_3 = 4$ .

Let  $C$  be a block code of length  $n$  over  $L$  which represents either an  $M$ -PSK or  $M$ -QASK signal set. If each component of codeword  $\mathbf{v}$  in  $C$  is mapped into its corresponding signal point in the 2-dimensional  $M$ -PSK or  $M$ -QASK signal set, we obtain a block  $M$ -PSK or  $M$ -QASK modulation code. The effective rate of this code is given by [9],

$$R[C] = \frac{1}{2n} \log_2 |C|. \quad (2.19)$$

### 3. Multi-Stage Decoding

Let  $L$  be the labeling for a  $h$ -dimensional signal set  $S$  with  $2^h$  signal points. Let  $C \triangleq C_1 C_2 \cdots C_m$  be an  $\ell$ -level code of length  $n$  over  $L$  with  $m$  component codes where  $C_i$  is a code over the sub-labeling  $L^{(i)}$ . In multi-stage decoding of  $C$ , component codes are decoded sequentially one at a time, stage by stage. The decoded information at each stage is passed to the next stage. The decoding process begins with the first component code  $C_1$  and ends at the last component code  $C_m$ . The decoding of a component code at each stage can be either soft-decision or hard-decision, maximum likelihood or bounded-distance decoding. As a result, there are four types of multi-stage decoding:

- (i) Multi-stage Soft-decision Maximum Likelihood Decoding—each stage of decoding is a soft-decision maximum likelihood decoding;
- (ii) Multi-stage Hard-decision Maximum Likelihood Decoding—each stage of decoding is a hard-decision maximum likelihood decoding;
- (iii) Multi-stage Bounded-distance Decoding—each stage of decoding is a bounded-distance decoding based on a certain distance measure, e.g., Hamming distance; and



(iv) Hybrid Multi-stage Decoding—mixed types of decoding are used among the stages.

In the following, we first describe a multi-stage decoding procedure and then formulate various types of stage decoding.

### A Multi-stage Decoding Procedure

Suppose a codeword in  $C$  is transmitted and  $\mathbf{z} = (z_1, z_2, \dots, z_n)$  is the received sequence at the output of the demodulator, where  $z_j$  is an  $h$ -tuple of real numbers. At the  $i$ -th stage of decoding with  $1 \leq i \leq m$ , the following process is carried out:

For  $2 \leq i \leq m$  and  $1 \leq j < i$ , let  $\mathbf{v}_D^{(j)}$  be the decoded codeword at the  $j$ -th stage decoding for  $C_j$ . Based on  $\mathbf{z}, \mathbf{v}_D^{(1)}, \dots, \mathbf{v}_D^{(i-1)}$  ( $\mathbf{z}$  for  $i=1$ ), the decoder performs a decoding procedure for  $C_i$  which is to be discussed later. Different kinds of decoding procedure may be used at different stages. If the decoding is successful, the decoder puts out a decoded codeword  $\mathbf{v}_D^{(i)}$  which is in  $C_i$ . Otherwise, stop the overall decoding and report that an uncorrectable error has been detected (this is a decoding failure).  $\Delta\Delta$

If every stage decoding is successful, then the decoded codeword  $\mathbf{v}_D$  in  $C$  is given by

$$\mathbf{v}_D = \mathbf{v}_D^{(1)} \mathbf{v}_D^{(2)} \dots \mathbf{v}_D^{(m)}. \quad (3.1)$$

Otherwise, an uncorrectable error has been detected and the decoder raises a flag.

Now we consider the decoding procedure at each stage. For  $1 \leq i \leq m$  and  $\mathbf{v}^{(1)} \mathbf{v}^{(2)} \dots \mathbf{v}^{(i-1)}$  in  $C_1 C_2 \dots C_{i-1}$  (the null string  $\lambda$  for  $i = 1$ ), let  $\overline{C}_i[\mathbf{v}^{(1)} \mathbf{v}^{(2)} \dots \mathbf{v}^{(i-1)}]$  be defined as the following set of vectors over  $L$ :

$$\begin{aligned} \overline{C}_i[\mathbf{v}^{(1)} \mathbf{v}^{(2)} \dots \mathbf{v}^{(i-1)}] &\triangleq \{ \mathbf{v}^{(1)} \mathbf{v}^{(2)} \dots \mathbf{v}^{(i-1)} \mathbf{v}^{(i)} \dots \mathbf{v}^{(m)} : \\ &\mathbf{v}^{(i)} \in C_i \text{ and } \mathbf{v}^{(j)} \in \{L^{(j)}\}^n \text{ for } i < j \leq m \}. \end{aligned} \quad (3.2)$$

It follows from (2.1) (2.2) (2.8) and (3.2) that for any distance measure  $d$ , the minimum distance of  $\overline{C}_i[\mathbf{v}^{(1)} \mathbf{v}^{(2)} \dots \mathbf{v}^{(i-1)}]$  is lower bounded as follows:

$$D[d, \overline{C}_i[\mathbf{v}^{(1)} \mathbf{v}^{(2)} \dots \mathbf{v}^{(i-1)}]] \geq D[d^{(i)}, C_i]. \quad (3.3)$$

where the equality holds if (2.10) and (2.11) are satisfied. In the following, we will show that the  $i$ -th stage decoding is a decoding procedure for  $\overline{C}_i[\mathbf{v}^{(1)} \mathbf{v}^{(2)} \dots \mathbf{v}^{(i-1)}]$ .

Hereafter we consider a memoryless additive channel and assume that every codeword of  $C$  is equally likely to be transmitted. For  $v \in L$ , let  $s(v)$  denote the signal point in  $R^h$

represented by  $v$ , where  $R^j$  denotes the set of all  $j$ -tuples of real numbers. For  $v \in L_{2^{\ell}\text{-PSK}}$ ,  $s(v)$  is given by  $(\cos(2^{1-\ell}\pi I(v)), \sin(2^{1-\ell}\pi I(v)))$ . For an  $n$ -tuple  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  over  $L$ , let  $s(\mathbf{v})$  denote the  $n$ -tuple  $(s(v_1), s(v_2), \dots, s(v_n))$  over  $S$ . For  $v \in L$  and  $z \in R^h$ , let  $pr(z|v)$  be the conditional probability that  $z$  is received given that the elementary signal represented by  $v$  is sent. Instead of  $pr(z|v)$ , we use a norm  $\|z - s(v)\|$  such that

$$\ln pr(z|v) = \gamma \|z - s(v)\|^2, \quad (3.4)$$

where  $\gamma$  is a negative constant real number and  $\ln$  denotes the natural logarithm. For an AWGN channel, we use the Euclidean distance in  $R^h$  as the norm. When  $L_{M\text{-PSK}}$  is considered, the channel is assumed to be an AWGN channel.

For an  $n$ -tuple  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  over  $R^h$ , let  $\|\mathbf{x}\|^2$  denote  $\sum_{j=1}^n \|x_j\|^2$ . Hereafter we take the distance measure  $d$  such that for  $u$  and  $v$  in  $L$ ,

(DM)

$$d(u, v) = \|s(u) - s(v)\|^2. \quad (3.5)$$

### Soft-Decision Maximum Likelihood Decoding at the $i$ -th Stage

Now we present a soft-decision maximum likelihood decoding at the  $i$ -th stage. The decoding is carried out as follows:

Let  $\mathbf{z} = (z_1, z_2, \dots, z_n)$  be the received vector at the output of the demodulator. Find a codeword  $\mathbf{v}$  in  $\overline{C}_i[\mathbf{v}^{(1)}\mathbf{v}^{(2)} \dots \mathbf{v}^{(i-1)}]$  for which the norm  $\|\mathbf{z} - s(\mathbf{v})\|^2$  is minimized. Then  $i$ -th component  $n$ -tuple of  $\mathbf{v}$  is the decoded codeword  $\mathbf{v}_D^{(i)} \in C_i$  for the  $i$ -th decoding stage.  $\triangle\triangle$

To carry out the above soft-decision maximum likelihood decoding at the  $i$ -th stage, it is desirable to choose the  $i$ -th component code  $C_i$  with a simple trellis diagram so that the Viterbi decoding algorithm can be used to reduce the number of computations. In this case, the metric for a branch labeled  $\mathbf{v}_q^{(i)} \in L^{(i)}$  corresponding to the  $q$ -th input symbol for  $1 \leq q \leq n$  is given by

$$\min_{u \in \{0,1\}^{l-j_{i+1}+1}} \|z_q - s(v_{Dq}^{(1)}v_{Dq}^{(2)} \dots v_{Dq}^{(i-1)}v_q^{(i)}u)\|^2, \quad (3.6)$$

where  $j_{i+1}$  is defined in (2.3).

Consider the soft-decision maximum likelihood decoding for the  $i$ -th component code of a multi-level  $M$ -PSK modulation code with  $M = 2^{\ell}$ . In this case,  $L = L_{M\text{-PSK}}$ . The signal

points in an  $M$ -PSK signal set  $S$  are labeled in such a way that the set of signal points whose labels have the same prefix  $a_1 a_2 \dots a_k$ , denoted  $Q(a_1 a_2 \dots a_k)$ , forms a  $M2^{-k}$ -PSK signal constellation (see Figure 1). This structure can be used to simplify the decoding. For  $z \in R^2$  and a binary string  $v$ , let  $T_v(z)$  denote the point in  $R^2$  which is obtained by rotating  $z$  around the origin by  $360I(v)/M$  degree clockwise, where  $I(v)$  is an integer defined in Section 2. For an  $n$ -tuple  $\mathbf{z} = (z_1, z_2, \dots, z_n)$  over  $R^2$  and an  $n$ -tuple  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  of binary strings, let  $T_{\mathbf{v}}(\mathbf{z})$  denote the  $n$ -tuple  $(T_{v_1}(z_1), T_{v_2}(z_2), \dots, T_{v_n}(z_n))$  over  $R^2$ . Let  $\overline{C}_i$  denote the following code over  $L_{M_i\text{-PSK}}$ :

$$\overline{C}_i \triangleq C_i \underbrace{V_n V_n \dots V_n}_{l-j_{i+1}+1}, \quad (3.7)$$

where  $M_i \triangleq M2^{-j_i+1}$  and  $V_n$  denotes the set of all binary  $n$ -tuples. Let  $\mathbf{z} = (z_1, z_2, \dots, z_n)$  be the output of the demodulator and  $\mathbf{v}_D^{(j)} = (v_{D1}^{(j)}, v_{D2}^{(j)}, \dots, v_{Dn}^{(j)})$  be the decoded codeword in  $C_j$  at the  $j$ -th stage for  $1 \leq j < i$ . Then it follows from the structure of  $M$ -PSK signal set that the  $i$ -th stage decoding based on  $\mathbf{z}$  and  $\mathbf{v}_D^{(1)}, \mathbf{v}_D^{(2)}, \dots, \mathbf{v}_D^{(i-1)}$  is reduced to decoding  $T_{\mathbf{v}_D^{(1)}, \mathbf{v}_D^{(2)}, \dots, \mathbf{v}_D^{(i-1)}}(\mathbf{z})$  for the  $M_i$ -PSK code  $\overline{C}_i$ .  $\Delta\Delta$

### Hard-Decision Decoding at the $i$ -th Stage

Suppose a signal point from a signal set  $S$  is transmitted. Let  $z \in R^h$  be the corresponding received point at the input of demodulator. The demodulator makes a hard decision (quantization) as follows:

For the given received point  $z$  and decoded sub-labels  $v^{(j)} \in L^{(j)}$  with  $1 \leq j < i$ , find the label  $v = v^{(1)}v^{(2)} \dots v^{(i-1)}v^{(i)} \dots v^{(m)}$  in  $L$  with  $v^{(1)}v^{(2)} \dots v^{(i-1)}$  as a prefix such that the norm  $\|z - s(v)\|$  is minimized. The  $i$ -th sub-label  $v^{(i)}$  of  $v$  is the hard-decision output of the demodulator.  $\Delta\Delta$

Let the hard-decision output of the demodulator at the  $i$ -th decoding stage be denoted by  $H_i(z, v^{(1)}v^{(2)} \dots v^{(i-1)})$  ( $H_1(z, \lambda)$  for  $i = 1$ ). Now we formulate a hard-decision maximum likelihood decoding at the  $i$ -th stage as follows:

For  $1 \leq j < i$ , let  $\mathbf{v}_D^{(j)} = (v_{D1}^{(j)}, v_{D2}^{(j)}, \dots, v_{Dn}^{(j)})$  be the decoded codeword in  $C_j$  at the  $j$ -th stage. For the output  $\mathbf{z} = (z_1, z_2, \dots, z_n)$  of the demodulator, let  $H_i(\mathbf{z}, \mathbf{v}_D^{(1)}, \mathbf{v}_D^{(2)}, \dots, \mathbf{v}_D^{(i-1)})$  denote the  $n$ -tuple over  $L^{(i)}$  whose  $q$ -th element is

$H_i(z_q, v_{Dq}^{(1)}, v_{Dq}^{(2)}, \dots, v_{Dq}^{(i-1)})$ . Decode

$$H_i(z, v_D^{(1)}, v_D^{(2)}, \dots, v_D^{(i-1)}) \quad (3.8)$$

into a codeword  $v_D^{(i)}$  in  $C_i$  such that  $d^{(i)}[H_i(z, v_D^{(1)}, v_D^{(2)}, \dots, v_D^{(i-1)}), v_D^{(i)}]$  is minimized where the distance measure  $d^{(i)}$  is defined by (2.8).  $\Delta\Delta$

### Bounded-Distance Decoding

Let  $\delta$  be a real number such that

$$0 \leq \delta < D[d^{(i)}, C_i]/\Delta_i. \quad (3.9)$$

where  $\Delta_i$  is the least real number such that for  $u, v$  and  $w$  in  $L^{(i)}$ ,

$$\Delta_i(d^{(i)}(u, w) + d^{(i)}(w, v)) \geq d^{(i)}(u, v).$$

The bounded-distance- $\delta$  decoding at the  $i$ -th stage is defined as follows:

For a received  $n$ -tuple  $H_i(z, v_D^{(1)}, v_D^{(2)}, \dots, v_D^{(i-1)})$  over  $L^{(i)}$ , if there exists a codeword  $v$  in  $C_i$  such that

$$d^{(i)}[H_i(z, v_D^{(1)}, v_D^{(2)}, \dots, v_D^{(i-1)}), v] \leq \delta/2$$

( $v$  is unique), then decode  $H_i(z, v_D^{(1)}, v_D^{(2)}, \dots, v_D^{(i-1)})$  into  $v$ . Otherwise, declare a decoding failure.  $\Delta\Delta$

In the case for which  $L^{(i)} = \{0, 1\}$ ,  $C_i$  is a binary code. It follows from (2.12) and (2.13) that the hard-decision maximum likelihood or bounded-distance decoding at the  $i$ -th stage can be done in terms of Hamming distance for which  $\Delta_i = 1$ .

Consider the case where  $L = L_{M\text{-PSK}}$  with  $M = 2^l$ . For  $z \in R^2$ , let  $J(z)$  denote the integer  $j$  such that  $0 \leq j < 2M$  and

$$\frac{\pi j}{M} < \varphi \leq \frac{\pi(j+1)}{M}, \quad (3.10)$$

where  $\varphi$  is the angle of the polar co-ordinates of  $z$ . Then it is readily seen [14] that the hard-decision  $H_i(z, v^{(1)}v^{(2)} \dots v^{(i-1)})$  can be determined by  $J(z)$ ,  $i$  and  $v^{(1)}v^{(2)} \dots v^{(i-1)}$ . Once  $J(z)$  is stored,  $z$  itself is unnecessary to be stored unless soft-decision decoding is used at a later stage. This reduces decoder complexity.

So far we have presented three basic types of decoding at a stage in a multi-stage decoding for a multi-level modulation code. Now we like to know under what conditions the decoding at a specific stage is correct. This is answered by Lemma 1 which follows from (2.1), (2.8), (3.3) and (3.5). This type of lemma was first given in [1] and then in [4] for some classes of multi-level codes.

**Lemma 1:** Suppose that (i) the sequence  $s(\mathbf{v})$  of elementary signals represented by a codeword  $\mathbf{v} = \mathbf{v}^{(1)}\mathbf{v}^{(2)} \dots \mathbf{v}^{(m)}$  in  $C$  is sent, (ii)  $\mathbf{z}$  is received, and (iii) for  $1 \leq i \leq m$ , the decoding at every stage prior to the  $i$ -th stage is correct. Then the decoding at the  $i$ -th stage is correct if one of the following conditions is met:

- (1) For soft-decision maximum likelihood decoding at the  $i$ -th stage,

$$\|\mathbf{z} - s(\mathbf{v})\|^2 < D[d^{(i)}, C_i]/4.$$

- (2) For hard-decision maximum likelihood decoding at the  $i$ -th stage,

$$\|\mathbf{z} - s(\mathbf{v})\|^2 \leq D[d^{(i)}, C_i]/(8\Delta_i).$$

- (3) For hard-decision bounded-distance- $\delta$  decoding at the  $i$ -th stage,

$$\|\mathbf{z} - s(\mathbf{v})\|^2 \leq \delta/8,$$

where  $0 \leq \delta < D[d^{(i)}, C_i]/\Delta_i$ .

$\Delta\Delta$

The three basic types of decoding presented above can be used at various stages to form various types of multi-stage decoding for a multi-level modulation code as we pointed out at the beginning of this section. A drawback of a multi-stage decoding is the error propagation effect caused by passing incorrectly decoded information from one stage to the next stage. As a result, the multi-stage soft-decision maximum likelihood decoding is not optimum even though the decoding at each stage is optimum. It is a suboptimum decoding. However, the error propagation effect can be made negligibly small, if the first few component codes (mostly the first component code) are powerful and decoded with the soft-decision maximum likelihood decoding. Based on our computations in next section, we find that the difference in performance between the suboptimum multi-stage decoding and the optimum single-stage decoding of a multi-level modulation is very small, only a fraction of a dB loss in coding gain. However the multi-stage decoding reduces the decoder complexity tremendously.

#### 4. Performance Analysis for Multi-Stage Decoding

In this section, error performance of the multi-stage decoding for multi-level modulation codes is analyzed for a memoryless additive channel, e.g., an AWGN channel. For an  $\ell$ -level code  $C = C_1 C_2 \dots C_m$  of length  $n$  with  $m$  component codes over  $L$ , we consider the  $i$ -th stage decoding with  $1 \leq i \leq m$ . For a codeword  $\mathbf{u}$  of  $C_i$ , let  $P_e^{(i)}(\mathbf{u})$  be the probability that every stage decoding prior to the  $i$ -th stage is correct but the  $i$ -th stage decoding is erroneous for

a received block when  $\mathbf{u}$  is transmitted, and let  $P_{df}^{(i)}(\mathbf{u})$  be the probability that every stage decoding prior to the  $i$ -th stage is correct but a decoding failure occurs at the  $i$ -th stage for a received block when  $\mathbf{u}$  is transmitted. We assume that every codeword is equally likely to be transmitted. Let  $p_e^{(i)}$  and  $p_{df}^{(i)}$  denote the average of  $P_e^{(i)}(\mathbf{u})$  and that of  $P_{df}^{(i)}(\mathbf{u})$  over all codewords, respectively. Let  $P_{ic}^{(i)}(\mathbf{u})$  and  $p_{ic}^{(i)}$  be defined as

$$P_{ic}^{(i)}(\mathbf{u}) \triangleq P_e^{(i)}(\mathbf{u}) + P_{df}^{(i)}(\mathbf{u}),$$

$$p_{ic}^{(i)} \triangleq p_e^{(i)} + p_{df}^{(i)}.$$

Let  $p_e$  denote the probability that a received block is erroneously decoded by the overall multi-stage decoder, and let  $p_{ic}$  denote the total probability that the overall multi-stage decoder fails in decoding a received block correctly (if there is no decoding failure, the former is equal to the latter). Then we readily see that

$$p_e \leq \sum_{i=1}^m p_e^{(i)}, \quad (4.1)$$

$$p_{ic} = \sum_{i=1}^m p_{ic}^{(i)}. \quad (4.2)$$

Clearly the probability  $p_{ic}$  can serve as an upper bound on the probability  $p_e$ .

First we present a sufficient condition for a code over  $L_{M\text{-PSK}}$  that  $p_e^{(i)} = P_e^{(i)}(\mathbf{u})$  and  $p_{df}^{(i)} = P_{df}^{(i)}(\mathbf{u})$  for any codeword  $\mathbf{u}$ , where  $M = 2^{\ell}$ . A code over  $L_{M\text{-PSK}}$  is said to be closed under component-wise modulo- $M$  addition, if and only if for any codewords  $\mathbf{u} = (u_1, u_2, \dots, u_n)$  and  $\mathbf{v} = (v_1, v_2, \dots, v_n)$ , there is a codeword  $\mathbf{v}' = (v'_1, v'_2, \dots, v'_n)$  such that for  $1 \leq j \leq n$

$$I(v'_j) = I(u_j) + I(v_j) \quad (\text{modulo } M). \quad (4.3)$$

Then the following lemma holds.

**Lemma 2:** Suppose that  $C = C_1 C_2 \dots C_m$  is a multi-level code over  $L_{M\text{-PSK}}$ . For  $1 \leq i \leq m$ , if (1) the component code  $C_i$  considered as a code over  $L_{2^{\ell^{(i)}}\text{-PSK}}$  is closed under component-wise modulo- $2^{\ell^{(i)}}$  addition where  $\ell^{(i)} = j_{i+1} - j_i$ , and (2) the soft-decision or hard-decision maximum likelihood decoding or the hard-decision bounded-distance decoding is used at the  $i$ -th stage, then it holds that for any codeword  $\mathbf{u}$  in  $C_i$ ,

$$p_e^{(i)} = P_e^{(i)}(\mathbf{u}), \quad (4.4)$$

$$p_{df}^{(i)} = P_{df}^{(i)}(\mathbf{u}). \quad (4.5)$$

**Proof:** It is easy to show that  $C_i$  is closed under component-wise modulo- $2^{\ell^{(i)}}$  addition if and only if  $\overline{C}_i$  defined by (3.7) is closed under component-wise modulo- $M_i$  addition. Then this lemma follows from (2.16).  $\triangle\triangle$

It follows from Lemma 2 that if  $\ell^{(i)} = 1$  and  $C_i$  is a binary linear code, then (4.4) and (4.5) are satisfied.

Next we evaluate  $P_e^{(i)}(\mathbf{0})$  and  $P_{df}^{(i)}(\mathbf{0})$  for an  $\ell$ -level code over  $L$ . For simplicity, we will consider the case where  $\ell^{(i)} = 1$ . Suppose that the all-zero  $n$ -tuple  $\mathbf{0}$  over  $L$  is sent and every stage decoding prior to the  $i$ -th stage is correct. Let  $\mathbf{z} = (z_1, z_2, \dots, z_n)$  be the output of the demodulator.

### Soft-Decision Maximum Likelihood Decoding

Consider the soft-decision maximum likelihood decoding at the  $i$ -th stage. In this case, there is no decoding failure, that is,  $p_{df}^{(i)} = P_{df}^{(i)}(\mathbf{0}) = 0$ . Let  $\mathbf{0}^j$  denote the  $n$ -tuple  $(0^j, 0^j, \dots, 0^j)$ , where  $0^j$  denotes the string of  $j$  zeros. For  $1 \leq i \leq m$ , let  $U_i$  denote  $L^{(i+1)}L^{(i+2)} \dots L^{(m)} (= \{0, 1\}^{\ell-j_{i+1}+1})$  and  $U_i^n$  denote the set of all  $n$ -tuples over  $U_i$ . For  $v \in L^{(i)} (= \{0, 1\})$  and  $u \in U_i$ , let  $s_i(vu)$  denote  $s(\mathbf{0}^{i-1}vu)$ , where  $s(u')$  is the signal point represented by  $u' \in L$ . If  $L = L_{M\text{-PSK}}$ , then  $s_i(u)$  is the  $M_i$ -PSK signal point represented by  $u \in L_{M\text{-PSK}}$ . For an  $n$ -tuple  $\mathbf{u} = (u_1, u_2, \dots, u_n)$  over  $L^{(i)}U_i$  (the concatenation of  $L^{(i)}$  and  $U_i$ ), let  $s_i(\mathbf{u})$  denote  $(s_i(u_1), s_i(u_2), \dots, s_i(u_n))$ . Then the decoding at the  $i$ -th stage is correct if for any nonzero codeword  $\mathbf{v}^{(i)} = (v_1^{(i)}, v_2^{(i)}, \dots, v_n^{(i)})$  in  $C_i$ ,

$$\min_{\mathbf{u} \in U_i^n} \|\mathbf{z} - s_i(\mathbf{v}^{(i)}\mathbf{u})\|^2 > \min_{\mathbf{u} \in U_i^n} \|\mathbf{z} - s_i(\mathbf{0}\mathbf{u})\|^2. \quad (4.6)$$

This inequality can be rearranged as follows:

$$\sum_{j \text{ such that } v_j^{(i)}=1} \left\{ \min_{u \in U_i} \|z_j - s_i(1u)\|^2 - \min_{u \in U_i} \|z_j - s_i(0u)\|^2 \right\} > 0. \quad (4.7)$$

A codeword  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  in  $C_i$  is said to be "nondecomposable" if  $\mathbf{v}$  cannot be expressed as a component-wise integer sum of two or more nonzero codewords in  $C_i$ . It is readily seen that only nondecomposable codewords of  $C_i$  need to be considered in (4.6) or (4.7). For a positive integer  $w$ , let  $q_i(w)$  denote the probability that the sum of  $w$  independent identically distributed random variables

$$\min_{u \in U_i} \|z - s_i(1u)\|^2 - \min_{u \in U_i} \|z - s_i(0u)\|^2 \quad (4.8)$$

is not positive, where  $z$  is a random variable with conditional density function  $p(z|0^t)$  and  $p(\cdot|\cdot)$  is the channel symbol transition probability. From (4.7) and (4.8), the probability that the decoded codeword at the  $i$ -th stage is a specific nonzero codeword of weight  $w$  is upper bounded by  $q_i(w)$ . For a positive integer  $w$ , let  $A'_w$  denote the number of nondecomposable codewords with Hamming weight  $w$  in  $C_i$ . Then,  $P_{ic}^{(i)}(0)$  is upper bounded as follows:

$$P_{ic}^{(i)}(0) \leq \sum_{w=1}^n A'_w q_i(w). \quad (4.9)$$

To compute the upper bound on  $P_{ic}^{(i)}(0)$  given by (4.9), we need to evaluate  $q_i(w)$  for  $1 \leq w \leq n$ . The parameter  $q_i(w)$  can be evaluated in the following ways. Let  $\underline{E}$  denote the infimum with respect to  $t$  of the moment generating function  $E(t)$  of the random variable (4.8). Then the following upper bound on  $q_i(w)$  holds:

$$q_i(w) \leq \underline{E}^w. \quad (4.10)$$

This upper bound [16] is useful for relatively large  $w$ . For  $L_{M\text{-PSK}}$  with  $M = 2^\ell$  and an AWGN channel, let  $z = (x, y) \in R^2$ , and  $M_i \triangleq 2^{\ell-j_{i+1}+1}$  with  $1 \leq i \leq m$ . Then,  $\min_{u \in U_i} \|z - s_i(1u)\|^2 - \min_{u \in U_i} \|z - s_i(0u)\|^2$ , denoted  $\xi$ , is lower bounded as follows:

(1) If  $y \geq 0$ , then

$$\xi \geq \|z - s_i(10^{\ell-j_{i+1}+1})\|^2 - \|z - s_i(0^{\ell-j_{i+1}+2})\|^2, \quad (4.11)$$

(2) if  $y \leq 0$ , then

$$\xi \geq \|z - s_i(1^{\ell-j_{i+1}+1})\|^2 - \|z - s_i(0^{\ell-j_{i+1}+2})\|^2. \quad (4.12)$$

A proof of the above bounds is given in Appendix A. It follows from (4.11) and (4.12) that  $\xi$  is lower bounded by

$$2x(1 - \cos \frac{2\pi}{M_i}) - 2|y| \sin \frac{2\pi}{M_i}. \quad (4.13)$$

Let  $q'_i(w)$  be the probability that the sum of  $w$  independent random variables (4.13) is not positive. Then

$$q_i(w) \leq q'_i(w). \quad (4.14)$$

It is easier to evaluate  $q'_i(w)$  than  $q_i(w)$ . The following upper bound on  $q'_i(w)$  holds (see Appendix B for the proof).

$$q'_i(w) \leq 2^w Q\left(\sqrt{\frac{wd_{j_i}\rho}{2}}\right), \quad \text{for } i < m \quad (4.15)$$



where  $d_{j_i}$  is defined by (2.18),  $\rho = 2R[C]E_b/N_0$ , and

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt. \quad (4.16)$$

If  $i = m$ , then  $M_i = 2$  and

$$q_m(w) = Q\left(\sqrt{\frac{wd_t\rho}{2}}\right). \quad (4.17)$$

### Hard-Decision Bounded-Distance Decoding

We consider the hard-decision bounded-distance- $\delta$  decoding at the  $i$ -th stage, where

$$\delta = 2d_{j_i}t, \quad (4.18)$$

$$2t + 1 \leq \delta_i, \quad (4.19)$$

where  $d_{j_i}$  is the  $j_i$ -th distance parameter of  $L$  defined by (2.14) and  $\delta_i$  is the minimum Hamming distance of  $C_i$ . To derive upper bounds on  $P_e^{(i)}(0)$  and  $P_{df}^{(i)}(0)$ , we consider the following decoding problem: Decode

$$H_i(\mathbf{z}, \mathbf{0}^{j_i-1}) \quad (4.20)$$

with bounded-distance- $\delta$  decoding, where  $H_i$  is defined in Section 3. Note that for  $1 \leq q \leq n$ , the  $q$ -th bit of  $H_i(\mathbf{z}, \mathbf{0}^{j_i-1})$  is  $H_i(z_q, \mathbf{0}^{j_i-1})$ . The probability that  $H_i(z_q, \mathbf{0}^{j_i-1}) = 1$  is  $q_i(1)$ . Let  $\mathbf{P}_c^{(i)}$ ,  $\mathbf{P}_e^{(i)}$  and  $\mathbf{P}_{df}^{(i)}$  be the probabilities of a correct decoding, an erroneous decoding and a decoding failure of the above decoding problem, respectively. We assume that  $C_i$  is linear. Then it holds [21] that

$$\mathbf{P}_c^{(i)} = \sum_{j \leq t} \binom{n}{j} q_i(1)^j (1 - q_i(1))^{n-j}, \quad (4.21)$$

$$\mathbf{P}_e^{(i)} = \sum_{w=\delta_i}^n A_w \sum_{w-h+j \leq t} \binom{w}{h} \binom{n-w}{j} q_i(1)^{h+j} (1 - q_i(1))^{n-h-j}, \quad (4.22)$$

$$\mathbf{P}_{df}^{(i)} = 1 - \mathbf{P}_c^{(i)} - \mathbf{P}_e^{(i)}. \quad (4.23)$$

In the derivation of the above formulas, it is disregarded whether for a given  $\mathbf{z}$ , the decoded codeword at the  $j$ -th stage with  $1 \leq j < i$  is  $\mathbf{0}^{(j)}$  or not. Consequently, the following inequalities hold:

$$P_e^{(i)}(0) \leq \mathbf{P}_e^{(i)}, \quad (4.24)$$

$$P_{df}^{(i)}(0) \leq \mathbf{P}_{df}^{(i)}, \quad (4.25)$$

$$P_{ic}^{(i)}(0) = 1 - \mathbf{P}_c^{(i)}. \quad (4.26)$$

The equalities in (4.24) and (4.25) hold for  $i = 1$ . For the case where hard-decision bounded-distance decoding is used at each previous stage, explicit formulas for  $p_e^{(i)}$ ,  $p_{ic}^{(i)}$  and  $p_{df}^{(i)}$  have been derived in [14, 15] and that for  $p_{ic}$  is shown in Appendix C.

In Figures 2 to 6, the error performance of various types of multi-stage decoding for several block multi-level modulation codes listed in Table 1 are shown. The channel is assumed to be an AWGN channel. In Table 1, the following notations are used:

- (1)  $P_n$  denotes the binary  $(n, n-1)$  linear code which consists of all the even-weight binary  $n$ -tuples.
- (2)  $RM_{i,j}$  denotes the  $j$ -th order Reed-Muller code of length  $n = 2^i$ .
- (3)  $BCH_{i,d}$  denotes the binary primitive BCH code of length  $2^i - 1$  and designed distance  $d$ .
- (4) For a code  $C$ ,  $ex-C$  denotes the extended code of  $C$  by adding an overall parity bit.
- (5)  $D_i$  denotes  $D[d^{(i)}, C_i]$  and  $d_i$  denotes the  $i$ -th distance parameter.
- (6) In the column of decoding type, for example,  $(3s)$  means that  $m = 1$ ,  $\ell^{(1)} = \ell = 3$  and single stage soft-decision maximum likelihood decoding is used.  $(s, s, s)$  means that  $m = 3$ ,  $\ell^{(1)} = \ell^{(2)} = \ell^{(3)}$  and soft-decision maximum likelihood decoding is used at each of the three decoding stages;  $(s, h_{t_2}, s)$  means that  $m = 3$ ,  $\ell^{(1)} = \ell^{(2)} = \ell^{(3)} = 1$  and soft-decision maximum likelihood decoding at the first and third stages, hard-decision bounded-distance- $(2t_2 + 1)$  decoding at the second stage are used. " $h_0$ " indicates no decoding.
- (7) In the column of complexity of trellis diagram, the  $i$ -th number means the number of states of a 4-section trellis diagram which is available at the  $i$ -th stage soft-decision maximum likelihood decoding. For a case where hard-decision bounded-distance decoding is used at the stage, "-" is marked.

In Figures 2 to 6, the following notations are used.

- (1)  $p_{ic}$  [decoding type] denotes the probability that the multi-stage decoder specified by the decoding type fails in decoding a received block correctly.
- (2)  $p_{ic,s}$  [decoding type] (or  $P_{ic,s}(0)$  [decoding type]) denotes the simulation result on  $p_{ic}$  [decoding type] (or  $P_{ic}(0)$  [decoding type]).
- (3)  $\bar{p}_{ic}$ [decoding type] denotes an upper bound on  $p_{ic}$  [decoding type] derived from (4.2) by replacing  $p_{ic}^{(i)}$  with its upper bound  $\bar{p}_{ic}^{(i)}$ . If the soft-decision maximum likelihood decoding

is used at the  $i$ -th stage,  $\bar{p}_{ic}^{(i)}$  is given by the right-hand side of (4.9) where  $q_i(w)$  is upper bounded by the minimum of the right-hand sides of (4.10) and (4.15) or (4.17). If the hard-decision bounded-distance decoding is used at the  $i$ -th stage,  $\bar{p}_{ic}^{(i)}$  is given by (4.26).  $\bar{P}_{ic}(\mathbf{0})[(3s)]$  denotes an upper bound on  $P_{ic}(\mathbf{0})[(3s)]$  given by (4.11) in [7].

- (4) For  $1 \leq \delta \leq n$ , let  $\bar{p}_{ic,s}^{(i)}(\delta)$  denote the value of the right-hand side of (4.9) where  $q_i(w)$  with  $\delta_i \leq w \leq \delta$  is evaluated by simulation results on (4.13) and  $q_i(w)$  with  $\delta < w \leq n$  is evaluated by the minimum of the right-hand sides of (4.10) and (4.15) (or (4.17)). In Figures 2 to 4,  $\bar{p}_{ic}[\text{decoding type}]$  denotes  $\bar{p}_{ic,s}^{(1)}(\delta_1) + \bar{p}_{ic,s}^{(2)}(\delta_2) + \bar{p}_{ic}^{(3)}$ .
- (5)  $\bar{p}_e$  [decoding type] denotes an upper bound on the probability that a received block is decoded erroneously by the hard-decision bounded-distance multi-stage decoding specified by the decoding type. The upper bound is computed by a formula in [14].

In these figures, the error performances are compared with those of some uncoded reference modulation systems for transmitting the same (or almost the same) number of information bits.

Figure 2 shows the error performance of the basic 3-level block 8-PSK modulation code  $C_1 = \text{RM}_{5,1}\text{RM}_{5,3}P_{32}$  with various types of decoding: single-stage soft-decision maximum likelihood decoding (optimal), 3-stage soft-decision maximum likelihood decoding (suboptimal), and 3-stage hard-decision bounded-distance decoding. Code  $C_1$  has effective rate almost equal to one, minimum squared Euclidean distance 8, and a 4-section trellis diagram with 512 states. It achieves 6 dB asymptotic coding gain over the uncoded QPSK system with the soft-decision maximum likelihood decoding. To carry out the one-stage optimal decoding with the Viterbi algorithm, a decoder of 512 states is needed, which is quite complex. From Figure 2, we see that, with the one-stage optimal decoding, the real coding gain of  $C_1$  over the uncoded QPSK system at the block-error-rate of  $10^{-6}$  is 5 dB. We also see that the difference in error performance between the 3-stage soft-decision maximum likelihood decoding and the one-stage optimal decoding is quite small. At block-error-rate of  $10^{-4}$ , the difference is only 0.3 dB (based on simulation results). The difference should be less than 0.3 dB for lower block-error-rates. Even based on the upper bound for  $p_{ic}$ , we see that the difference is less than 0.5 dB for block-error-rates below  $10^{-6}$ . This says that with the suboptimal 3-stage soft-decision maximum likelihood decoding, the loss of coding gain compared to the one-stage optimal decoding

is small. With the 3-stage soft-decision maximum likelihood decoding using the Viterbi algorithm, three small Viterbi decoders are required, a 16-state Viterbi decoder at the first stage, a 16-state Viterbi decoder at the second stage, and a two-state Viterbi decoder at the third stage. The overall 3-stage decoder has a total of 34 states compared with 512 states for the single-stage decoder for the overall code  $C_1$ . We see that there is a tremendous reduction in decoding complexity by using multi-stage decoding. This big reduction in decoding complexity represents an **excellent trade-off** for the small loss in coding gain. From Figure 2, we also see that even the 3-stage hard-decision bounded-distance decoding of  $C_1$ , denoted  $(h_7, h_1, h_0)$ , achieves very good error performance compared with the single-stage optimal decoding. There is a loss of 2.2 dB at the block-error-rate of  $10^{-6}$ , but there is still 2.7 dB coding gain over the uncoded QPSK system. With the 3-stage hard-decision bounded-distance decoding, the decoding complexity is further reduced (note that the first-level and second-level component codes of  $C_1$  are Reed-Muller codes which are majority-logic decodable).

Figure 3 shows the error performance of the basic 3-level 8-PSK block modulation code  $C_2 = \text{RM}_{6,2}\text{RM}_{6,4}\text{P}_{64}$  with the 3-stage soft-decision maximum likelihood decoding and the 3-stage hard-decision bounded-distance decoding. This code has effective rate  $R[C_2] = 1.11$  and minimum squared Euclidean distance 8. It does have a 4-section trellis diagram but with a very large number of states. Decoding this code with the single-stage soft-decision maximum likelihood decoding using Viterbi algorithm is prohibitively complex. However, with 3-stage soft-decision maximum likelihood decoding, the code achieves a 4.4 dB coding gain over the uncoded QPSK system at the block-error-rate  $10^{-6}$  with a big reduction in decoding complexity. In fact, this coding gain is achieved with bandwidth **reduction**. With the 3-stage hard-decision bounded-distance decoding, denoted  $(h_7, h_1, h_0)$ , the code also achieves significant coding gain over the uncoded QPSK system with bandwidth reduction. There is a 2.2 dB loss in coding gain compared with the 3-stage soft-decision maximum likelihood decoding, however the decoding complexity is greatly reduced.

Figure 4 shows the error performance of the basic 3-level 8-PSK block modulation code  $C_3 = \text{RM}_{6,1}\text{RM}_{6,4}\text{P}_{64}$  with the 3-stage soft-decision maximum likelihood decoding. This code has effective rate almost equal to one and minimum squared Euclidean distance 8. The code has a 4-section trellis diagram with 2048 states. The single-stage soft-decision maximum likelihood decoding using Viterbi algorithm may be too complex to implement. With the 3-

stage soft-decision maximum likelihood decoding, the code achieves a 4.3 dB coding gain over the uncoded QPSK system at the block-error-rate  $10^{-6}$  with almost no bandwidth expansion. Using Viterbi decoding algorithm at each stage, the overall decoder consists of three Viterbi decoders, a 32-state Viterbi decoder at the first stage, a 32-state Viterbi decoder at the second stage, and a two-state Viterbi decoder at the third stage. The total number of states for the 3-stage decoder is 66 compared to 2048 states for the single-stage decoder for the code. Note that the code achieves a 6 dB asymptotic coding gain over uncoded QPSK system with the single-stage optimal decoding. Since the number of states of trellis diagram of the code is too big, simulation of the error performance of the code with the single-stage optimal decoding using Viterbi algorithm is very time consuming and hence is not being carried out. The real coding gain of the code over the uncoded QPSK system is unlikely more than 5 dB at block-error-rate of  $10^{-6}$  with the single-stage optimal decoding.

Figure 5 shows the error performance of the basic 3-level 8-PSK block modulation code  $C_4 = \text{RM}_{6,1}\text{ex-BCH}_{6,7}\text{P}_{64}$  with the hybrid 3-stage decoding in which the first stage and third stage decodings are soft-decision maximum likelihood decodings and the second stage decoding is hard-decision bounded-distance decoding. The second component code of  $C_4$  is an extended BCH code of length 64 which has a very complex trellis diagram. Therefore we choose hard-decision bounded-distance decoding for this component code. From Figure 5, we see that with the 3-stage hybrid decoding, the code achieves a 4.4 dB coding gain over the uncoded QPSK system at block-error-rate  $10^{-6}$  with 10% bandwidth expansion.

Figure 6 shows the error performance of two basic 3-level 8-PSK block modulation codes with 3-stage hard-decision bounded-distance decoding. Both codes consist of BCH codes as component codes and both achieve more than 4 dB coding gain over the uncoded QPSK system at block-error-rate  $10^{-6}$  without bandwidth expansion (code  $C_6$  needs a little bandwidth expansion).

## 5. Conclusion

In this paper we have investigated various types of multi-stage decoding for multi-level block modulation codes. Analysis of error performance for these decoding schemes has been carried out. Based on our computation and simulation results for some bandwidth efficient multi-level

block modulation codes, we have found that the multi-stage decoding provides an excellent trade-off between error performance and decoding complexity. The multi-stage soft-decision maximum likelihood decoding achieves an error performance close to that of the single-stage optimal decoding but with a great reduction in decoding complexity. We have also shown that hybrid multi-stage decoding should be used for those multi-level codes in which some component codes are too complex to decode with the soft-decision maximum likelihood decoding. Our conclusion is that multi-stage decoding of multi-level modulation codes offers the best of three worlds, spectral efficiency, error performance (or coding gain), and decoding complexity. Even though the various types of decoding are formulated for multi-level block modulation codes, they can be readily modified for multi-level trellis modulation codes.

## References

- [1] H. Imai and S. Hirakawa, "A New Multilevel Coding Method Using Error Correcting Codes," *IEEE Trans. on Information Theory*, Vol. IT-23, No. 3, pp. 371-376, May 1977.
- [2] V.V. Ginzburg, "Multidimensional Signals for a Continuous Channel," *Problemy pere-dachi Informatsii*, Vol. 20, No. 1, pp. 28-46, 1984.
- [3] S.I. Sayegh, "A Class of Optimum Block Codes in Signal Space," *IEEE Trans. on Com-munications*, Vol. COM-30, No. 10, pp. 1043-1045, October 1986.
- [4] R.M. Tanner, "Algebraic Construction of Large Euclidean Distance Combined Coding Modulation Systems," *Abstracts of Papers, 1986 IEEE International Symposium on In-formation Theory*, Ann Arbor, October 6-9, 1986.
- [5] G.J. Pottie and D.P. Taylor, "Multi-Level Channel Codes Based on Partitioning," *IEEE Trans. on Information Theory*, Vol. IT-35, No. 1, pp. 87-98, January 1989.
- [6] A.R. Calderbank, "Multi-Level Codes and Multi-Stage Decoding," *IEEE Trans. on Com-munications*, Vol. COM-37, No. 3, pp. 222-229, March 1989.
- [7] T. Kasami, T. Takata, T. Fujiwara and S. Lin, "On Multi-Level Block Modulation Codes," *IEEE Trans. on Information Theory*, Vol. IT-37, No. 4, July 1991.
- [8] T. Kasami, T. Takata, T. Fujiwara and S. Lin, "On Linear Structure and Phase Rotation Invariant Properties of Block  $2^L$ -PSK Modulation Codes," *IEEE Trans. on Information Theory*, Vol. IT-37, No. 1, pp. 164-167, January 1991.
- [9] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," *IEEE Trans. on Infor-mation Theory*, Vol. IT-28, No. 1, pp. 55-67, January 1982.
- [10] G. Ungerboeck, "Trellis-Coded Modulation with Redundant Signal Set, Part II: State of the Art," *IEEE Communications Magazine*, Vol. 25, No. 2, pp. 12-21, February 1987.
- [11] G.D. Forney, Jr., R.G. Gallager, C.R. Lang, F.M. Longstaff and S.U. Qureschi, "Efficient Modulation for Band-Limited Channels," *IEEE Journal of Select. Areas Communications*, Vol. SAC-2, No. 5, pp. 632-647, September 1984.
- [12] J. Leech and N.J.A. Sloane, "Sphere Packing and Error-Correcting Codes," *Can. J. Math.*, Vol. XXIII, No. 4, pp. 718-745, 1971.
- [13] T. Kasami, T. Takata, T. Fujiwara and S. Lin, "A Concatenated Coded Modulation Scheme for Error Control," *IEEE Trans. on Communications*, Vol. COM-38, No. 6, pp. 752-763, June 1990.
- [14] S. Ujita, "A Study on Multi-Stage Decoding of Block Modulation Codes," *Master's thesis, Graduate School of Osaka University*, February 1990.
- [15] M. Kataoka, S. Ujita, T. Takata, and T. Kasami, "Error Probability of Multi-stage Hard-decision Decoding for Block Modulation Codes," *the Proceedings of the 13th Symposium on Information Theory and Its Applications*, Vol. 1, pp. 197-202, Tatehina, Japan, Jan-uary 23-26, 1991, and also submitted to *Trans. of the Institute of Electronics, Information and Communication Engineers in Japan*, March 1991.

- [16] H. Chernoff, "A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations," *Ann. Math. Stat.*, Vol. 23, pp. 493-507, 1954.
- [17] A.R. Calderbank and N.J.A. Sloane, "New Trellis Codes Based on Lattices and Cosets," *IEEE Trans. on Information Theory*, Vol. IT-33, No. 2, pp. 177-195, March 1987.
- [18] T. Takata, S. Ujita, T. Fujiwara, T. Kasami and S. Lin, "Linear Structure and Error Performance Analysis of Block PSK Modulation Codes," *Transactions of IEICE of Japan*, Vol. J73-A, No. 2, pp. 314-321, February 1990.
- [19] S. Ujita, T. Takata, T. Fujiwara, T. Kasami and S. Lin, "A Multi-stage Decoding for Block Modulation Codes and Its Error Probability Analysis," *the Proceedings of the 12th Symposium on Information Theory and Its Applications*, Vol. 1, pp. 225-230, Inuyama, Japan, December 6-9, 1989.
- [20] G.D. Forney, Jr., "Coset Codes—Part I: Introduction and Geometrical Classification," *IEEE Trans. on Information Theory*, Vol. IT-34, No. 5, pp. 1123-1151, September 1988.
- [21] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, 1977.
- [22] G.D. Forney, Jr., "Coset Codes—Part II: Binary Lattices and Related Codes," *IEEE Trans. on Information Theory*, Vol. IT-34, No. 5, pp. 1152-1187, September 1988.
- [23] G.D. Forney, Jr., "A Bounded-Distance Decoding Algorithm for the Leech Lattice, with Generalizations," *IEEE Trans. on Information Theory*, Vol. IT-35, No. 4, pp. 906-909, July 1989.
- [24] J.C. Proakis, *Digital Communications*, McGraw-Hill, second edition, 1989.
- [25] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I, John Wiley & Sons, 1966.



## **ACKNOWLEDGMENT**

The authors would like to thank Mr. Yamashita for writing the simulation programs, and Mr. Kataoka for typing the manuscripts.

## Appendix A

### Proof of (4.11) and (4.12)

Let  $z \triangleq (x, y) \in R^2$  and  $\varphi \triangleq \tan^{-1}(y/x)$ . Let  $j$  be the nonnegative integer less than  $M_i$  such that

$$\frac{2\pi j}{M_i} \leq \varphi < \frac{2\pi(j+1)}{M_i}. \quad (\text{A.1})$$

For simplicity,  $s_i(u)$  with  $u \in L_{M_i\text{-PSK}}$  is denoted by  $\sigma(I(u))$ . Then,

$\xi (\triangleq \min_{u \in U_i} \|z - s_i(1u)\|^2 - \min_{u \in U_i} \|z - s_i(0u)\|^2)$  can be expressed as follows:

If  $j$  is even,

$$\xi = \|z - \sigma(j+1)\|^2 - \|z - \sigma(j)\|^2 \quad (\text{A.2})$$

and otherwise,

$$\xi = \|z - \sigma(j)\|^2 - \|z - \sigma(j+1)\|^2. \quad (\text{A.3})$$

By symmetry, it suffices to prove (4.11), and therefore,  $j$  is assumed to be less than  $M_i/2$ .

The left-hand side of (4.11), denoted  $\xi_0$ , can be expressed as

$$\xi_0 = \|z - \sigma(1)\|^2 - \|z - \sigma(0)\|^2. \quad (\text{A.4})$$

If  $j = 0$ , then  $\xi = \xi_0$ . We consider the remaining case where

$$1 \leq j < M_i/2. \quad (\text{A.5})$$

For  $z_1$  and  $z_2$  in  $R^2$ ,  $(z_1, z_2)$  denotes the inner product of  $z_1$  and  $z_2$ . It follows from (A.2) to (A.5) that

$$\begin{aligned} \xi_0 - \xi &= 2(z, \sigma(0) - \sigma(1) \pm (\sigma(j) - \sigma(j+1))) \\ &= 2\|z\| \left( \cos \varphi - \cos\left(\varphi - \frac{2\pi}{M_i}\right) \pm \left( \cos\left(\varphi - \frac{2\pi j}{M_i}\right) - \cos\left(\frac{2\pi(j+1)}{M_i} - \varphi\right) \right) \right) \\ &= -4\|z\| \sin \frac{\pi}{M_i} \left( \sin\left(\varphi - \frac{\pi}{M_i}\right) \pm \sin\left(\varphi - \frac{\pi(2j+1)}{M_i}\right) \right) \\ &= -8\|z\| \sin \frac{\pi}{M_i} \sin\left(\varphi - \frac{\pi(J+1)}{M_i}\right) \cos \frac{\pi j}{M_i} \end{aligned} \quad (\text{A.6})$$

$$\text{or} \quad -8\|z\| \sin \frac{\pi}{M_i} \cos\left(\varphi - \frac{\pi(J+1)}{M_i}\right) \sin \frac{\pi j}{M_i}. \quad (\text{A.7})$$

Inequalities (A.1) and (A.5) imply that

$$0 \leq \frac{\pi(j-1)}{M_i} \leq \varphi - \frac{\pi(j+1)}{M_i} < \frac{\pi j}{M_i} < \frac{\pi}{2}. \quad (\text{A.8})$$

It follows from (A.6), (A.7) and (A.8) that

$$\xi \geq \xi_0. \quad (\text{A.9})$$

$\Delta\Delta$

## Appendix B

### Proof of (4.15)

Note that  $q'_i(w)$  is the probability that the following inequality holds:

$$\sum_{j=1}^w (x_j - 1) \left(1 - \cos \frac{2\pi}{M_i}\right) - |y_j| \sin \frac{2\pi}{M_i} \leq -w \left(1 - \cos \frac{2\pi}{M_i}\right), \quad (\text{B.1})$$

where  $x_j$  and  $y_j$  with  $1 \leq j \leq w$  are independent Gaussian random variables with zero mean and variance  $\frac{1}{2\rho}$  where  $\rho = 2R[C]E_b/N_o$ . For a  $w$ -tuple  $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_w)$  over  $\{1, -1\}$ , let  $f_\epsilon$  denote the following random variable:

$$f_\epsilon \triangleq \sum_{j=1}^w (x_j - 1) \left(1 - \cos \frac{2\pi}{M_i}\right) + \epsilon_j y_j \sin \frac{2\pi}{M_i}. \quad (\text{B.2})$$

Then,  $f_\epsilon$  is a random variable with zero mean and variance  $wd_{j,i}/(2\rho)$ , and therefore, the probability that  $f_\epsilon \leq -w(1 - \cos \frac{2\pi}{M_i}) = -wd_{j,i}/2$  is given by

$$Q\left(\sqrt{\frac{wd_{j,i}\rho}{2}}\right). \quad (\text{B.3})$$

Assume that  $i < m$ . Then  $\sin \frac{2\pi}{M_i} > 0$ . Since the left-hand side of (B.1) is equal to  $\min_{\epsilon \in \{1, -1\}^w} f_\epsilon$ ,  $q'_i(w)$  is equal to the probability that at least one of  $f_\epsilon$  with  $\epsilon \in \{1, -1\}^w$  is less than or equal to  $-w(1 - \cos \frac{2\pi}{M_i})$ , which is upper bounded by  $2^w Q(\sqrt{\frac{wd_{j,i}\rho}{2}})$  [Bonferroni's inequality, 25].

△△

## Appendix C

For  $1 \leq i \leq \ell$ ,  $\beta = b_1 b_2 \cdots b_i \in \{0, 1\}^i$  and  $\gamma = c_1 c_2 \cdots c_{i-1} \in \{0, 1\}^{i-1}$ , let  $p_{\beta, \gamma}^{(i)}(u)$  denote the probability that  $H_h(z, c_1 c_2 \cdots c_{h-1}) = b_h$ , for  $1 \leq h \leq i$ . For simplicity, let  $p_{\beta, \gamma}^{(i)}$  denote  $p_{\beta, \gamma}^{(i)}(0^\ell)$ , where  $0^j$  denotes the string of  $j$  zeros.

For  $u \in L_{2^\ell \text{PSK}}$ , let  $\varphi$  denote the phase of the received  $z \in R^2$  when signal  $s(u)$  is sent. For  $0 \leq j \leq 2^{\ell+1} - 1$ , let  $q_j$  denote the probability that  $2^{-\ell}\pi j \leq \varphi - 2^{-\ell+1}\pi I(u) < 2^{-\ell}\pi(j+1)$ . For  $\ell = 3$ , the following equalities hold:

$$\begin{aligned}
 p_{0, \lambda}^{(1)}(u) &= q_0 + q_3 + q_4 + q_7 + q_8 + q_{11} + q_{12} + q_{15}, \\
 p_{1, \lambda}^{(1)}(u) &= q_1 + q_2 + q_5 + q_6 + q_9 + q_{10} + q_{13} + q_{14}, \\
 p_{00, 0}^{(2)}(u) &= q_0 + q_{15} + q_7 + q_8, \\
 p_{00, 1}^{(2)}(u) &= q_0 + q_3 + q_8 + q_{11}, \\
 &\vdots \\
 p_{11, 1}^{(2)}(u) &= q_5 + q_6 + q_{13} + q_{14}, \\
 p_{000, 00}^{(3)}(u) &= q_0 + q_{15}, \\
 p_{000, 01}^{(3)}(u) &= q_1 + q_{14}, \\
 &\vdots \\
 p_{111, 11}^{(3)}(u) &= q_{13} + q_{14}.
 \end{aligned} \tag{C.1}$$

It follows from the rotational symmetry of the signal points of  $L_{2^\ell \text{PSK}}$  it follows that for  $u = u_1 u_2 \cdots u_\ell \in L_{2^\ell \text{PSK}}$  and  $\beta = b_1 b_2 \cdots b_i \in \{0, 1\}^i$ ,

$$p_{b_1, \lambda}^{(1)}(u) = p_{b_1 \oplus u_1, \lambda}^{(1)}(0^\ell), \tag{C.2}$$

$$p_{\beta, u_1 u_2 \cdots u_{i-1}}^{(i)}(u) = p_{(b_1 \oplus u_1)(b_2 \oplus u_2) \cdots (b_i \oplus u_i), 0^{i-1}}^{(i)}(0^\ell), \tag{C.3}$$

where  $\oplus$  denotes modulo-2 addition.

Suppose that  $C \triangleq C^{(1)}C^{(2)} \cdots C^{(\ell)}$  is a basic  $\ell$ -level code of length  $n$  over  $L$  and that  $(2t_i + 1)$ -bounded-distance decoding is used at the  $i$ -th stage with  $1 \leq i \leq \ell$ . Let  $\mathbf{u} = \mathbf{u}^{(1)}\mathbf{u}^{(2)} \cdots \mathbf{u}^{(\ell)} = (u_1, u_2, \dots, u_n)$  be a codeword in  $C$ , where  $\mathbf{u}^{(h)} = (u_1^{(h)}, u_2^{(h)}, \dots, u_n^{(h)})$  for  $1 \leq h \leq \ell$  and  $u_j = u_j^{(1)}u_j^{(2)} \cdots u_j^{(\ell)}$  for  $1 \leq j \leq n$ . For  $1 \leq j \leq n$  and  $1 \leq h \leq \ell$ , let  $u_j^{(1, h)}$  and

$\mathbf{u}^{(1,h)}$  denote  $u_j^{(1)}u_j^{(2)}\dots u_j^{(h)}$  and  $\mathbf{u}^{(1)}\mathbf{u}^{(2)}\dots\mathbf{u}^{(h)}$ , respectively. Suppose that  $s(\mathbf{u})$  is sent and  $\mathbf{z} = (z_1, z_2, \dots, z_n)$  is received, where  $z_j \in R^2$ , and that for  $1 \leq h < \ell$ , the  $h$ -th stage decoding is successful (correct or erroneous). For  $1 \leq h \leq \ell$ , let  $\mathbf{v}^{(h)} = (v_1^{(h)}, v_2^{(h)}, \dots, v_n^{(h)})$  denote the decoded codeword in  $C^{(h)}$  at the  $h$ -th stage. Let  $\mathbf{v} \triangleq \mathbf{v}^{(1)}\mathbf{v}^{(2)}\dots\mathbf{v}^{(\ell)} \in C^{(1)}C^{(2)}\dots C^{(\ell)}$ ,  $v_j \triangleq v_j^{(1)}v_j^{(2)}\dots v_j^{(\ell)}$ ,  $v_j^{(1,h)} \triangleq v_j^{(1)}v_j^{(2)}\dots v_j^{(h)}$  and  $\mathbf{v}^{(1,h)} \triangleq \mathbf{v}^{(1)}\mathbf{v}^{(2)}\dots\mathbf{v}^{(h)}$  for  $1 \leq h \leq \ell$  and  $1 \leq j \leq n$ . For  $1 \leq h \leq \ell$ , let  $u_j^{(\ell,h)} \triangleq H_h(z_j, v_j^{(1)}v_j^{(2)}\dots v_j^{(h-1)})$  with  $1 \leq j \leq n$ ,  $\mathbf{u}^{(\ell,h)} \triangleq (u_1^{(\ell,h)}, u_2^{(\ell,h)}, \dots, u_n^{(\ell,h)})$  and  $\mathbf{u}' \triangleq \mathbf{u}^{(1)}\mathbf{u}^{(2)}\dots\mathbf{u}^{(\ell)} \triangleq (u'_1, u'_2, \dots, u'_n)$  where  $u'_j = u_j^{(1)}u_j^{(2)}\dots u_j^{(\ell)}$ . The condition that  $\mathbf{u}^{(\ell,h)}$  is decoded into  $\mathbf{v}^{(h)}$  at the  $h$ -th stage for  $1 \leq h \leq \ell$  is as follows:

$$d_H(\mathbf{u}^{(\ell,h)}, \mathbf{v}^{(h)}) \leq t_h, \quad \text{for } 1 \leq h \leq \ell, \quad (\text{C.4})$$

where  $d_H$  denotes the Hamming distance. For an  $n$ -tuple  $\mathbf{v} = \mathbf{v}^{(1)}\mathbf{v}^{(2)}\dots\mathbf{v}^{(\ell)}$  over  $L_{2^\ell, \text{PSK}}$ , let  $V_{t_1, t_2, \dots, t_\ell}^{(\ell)}(\mathbf{v})$  be defined as the set of  $n$ -tuples  $\mathbf{u}' = \mathbf{u}^{(1)}\mathbf{u}^{(2)}\dots\mathbf{u}^{(\ell)} = (u'_1, u'_2, \dots, u'_n)$ 's over  $L_{2^\ell, \text{PSK}}$  which satisfy (C.4). Let  $Q^{(\ell)}(\mathbf{u}, \mathbf{v})$  be the probability that when  $s(\mathbf{u})$  is sent, a received  $n$ -tuple over  $R^2$  is decoded into  $\mathbf{v}^{(h)}$  at the  $h$ -th stage for  $1 \leq h \leq \ell$ . Then it follows from the definition of  $p_{\beta, \gamma}^{(\ell)}(\mathbf{u})$  and (C.4) that

$$Q^{(\ell)}(\mathbf{u}, \mathbf{v}) = \sum_{(u'_1, u'_2, \dots, u'_n) \in V_{t_1, t_2, \dots, t_\ell}^{(\ell)}(\mathbf{v})} \prod_{j=1}^n p_{u'_j, v_j^{(1, \ell-1)}}^{(\ell)}(u_j). \quad (\text{C.5})$$

Suppose that for a codeword  $\mathbf{u} \in C$ ,  $s(\mathbf{u})$  is transmitted. Then, let  $p_c(\mathbf{u})$  denote the probability of correct decoding for all the  $\ell$  stages. It follows from (C.5) that

$$p_c(\mathbf{u}) = \sum_{(u'_1, u'_2, \dots, u'_n) \in V_{t_1, t_2, \dots, t_\ell}^{(\ell)}(\mathbf{u}^{(1, \ell)})} \prod_{j=1}^n p_{u'_j, u_j^{(1, \ell-1)}}^{(\ell)}(u_j). \quad (\text{C.6})$$

For  $\beta = b_1b_2\dots b_n$  and  $\gamma = c_1c_2\dots c_n$  in  $\{0, 1\}^n$ , let  $\beta \oplus \gamma$  denote  $(b_1 \oplus c_1)(b_2 \oplus c_2)\dots(b_n \oplus c_n)$ , and for  $n$ -tuples  $\beta$  and  $\gamma$  over  $\{0, 1\}^n$ , let  $\beta \oplus \gamma \triangleq (\beta_1 \oplus \gamma_1, \beta_2 \oplus \gamma_2, \dots, \beta_n \oplus \gamma_n)$ . The following lemma holds.

**Lemma C.1 :** For  $L_{2^\ell, \text{PSK}}$  and  $\mathbf{u} \in C$ , it holds that if  $C$  contains the zero word  $\mathbf{0}$ , then

$$p_c(\mathbf{u}) = p_c(\mathbf{0}) \quad (\text{C.7})$$

(Proof) It follows from (C.2), (C.3) and (C.6) that for  $\mathbf{u} \in C$ ,

$$p_c(\mathbf{u}) = \sum_{(u'_1, u'_2, \dots, u'_n) \in V_{t_1, t_2, \dots, t_\ell}^{(\ell)}(\mathbf{u}^{(1, \ell)})} \prod_{j=1}^n p_{u'_j \oplus u_j^{(1, \ell)}, 0^{(\ell-1)}}^{(\ell)}(0^\ell)$$

$$\begin{aligned}
&= \sum_{(u'_1, u'_2, \dots, u'_n) \in V_{t_1, t_2, \dots, t_\ell}^{(\ell)}(\mathbf{0}^\ell)} \prod_{j=1}^n p_{u'_j, \mathbf{0}^{\ell-1}}^{(\ell)}(\mathbf{0}^\ell) \\
&= p_c(\mathbf{0}).
\end{aligned} \tag{C.8}$$

$\Delta\Delta$

This lemma implies that for  $L_{2^\ell, \text{PSK}}$ ,

$$p_{ic} = 1 - p_c(\mathbf{0}). \tag{C.9}$$

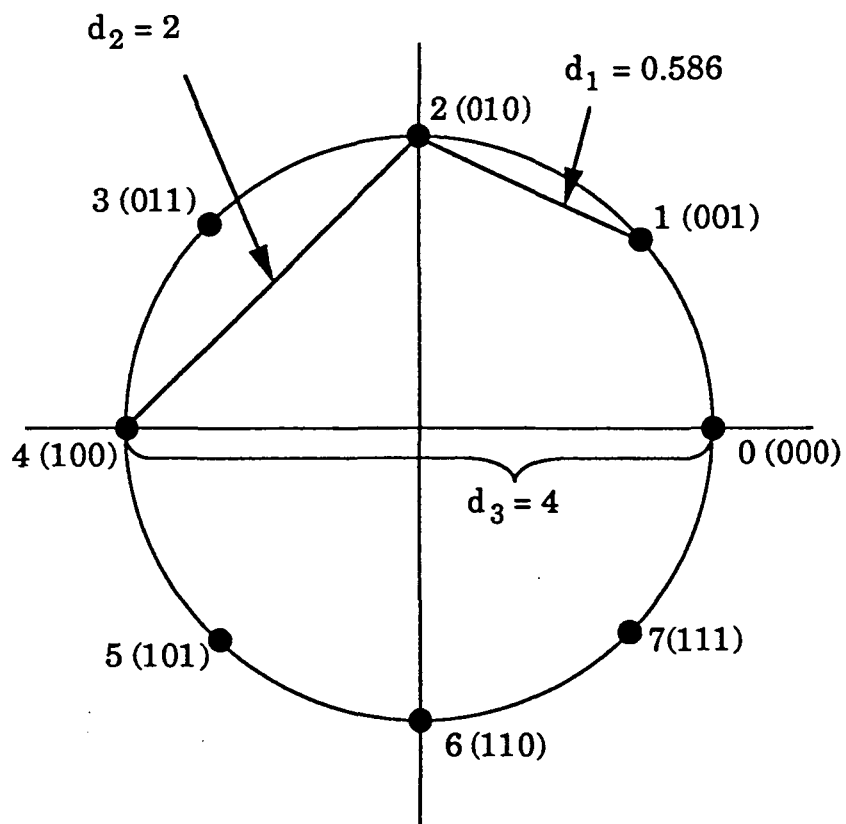


Fig. 1. 8-PSK signal set.



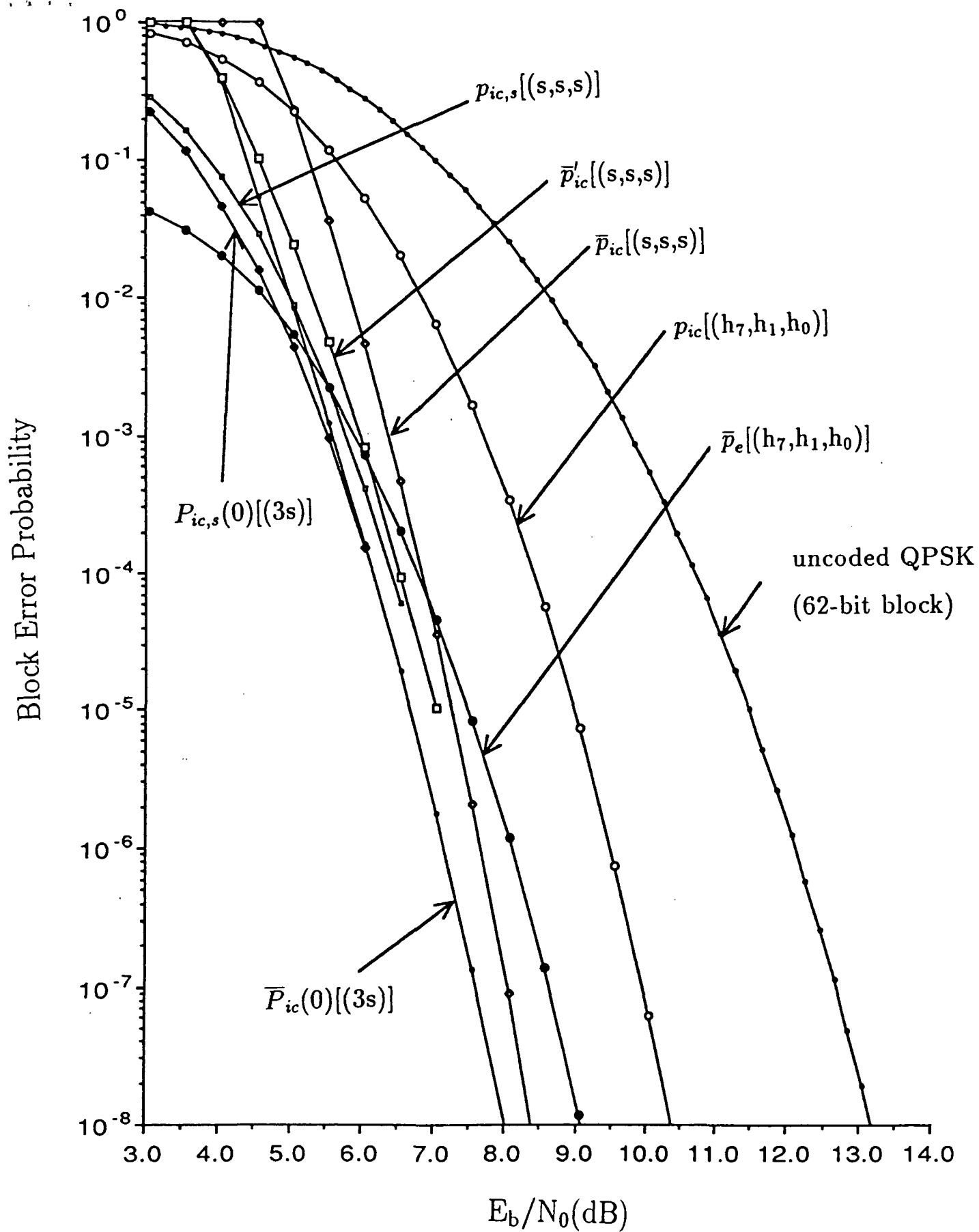


Fig.2. Error performance of the 8-PSK code  $RM_{5,1}RM_{5,3}P_{32}$ .

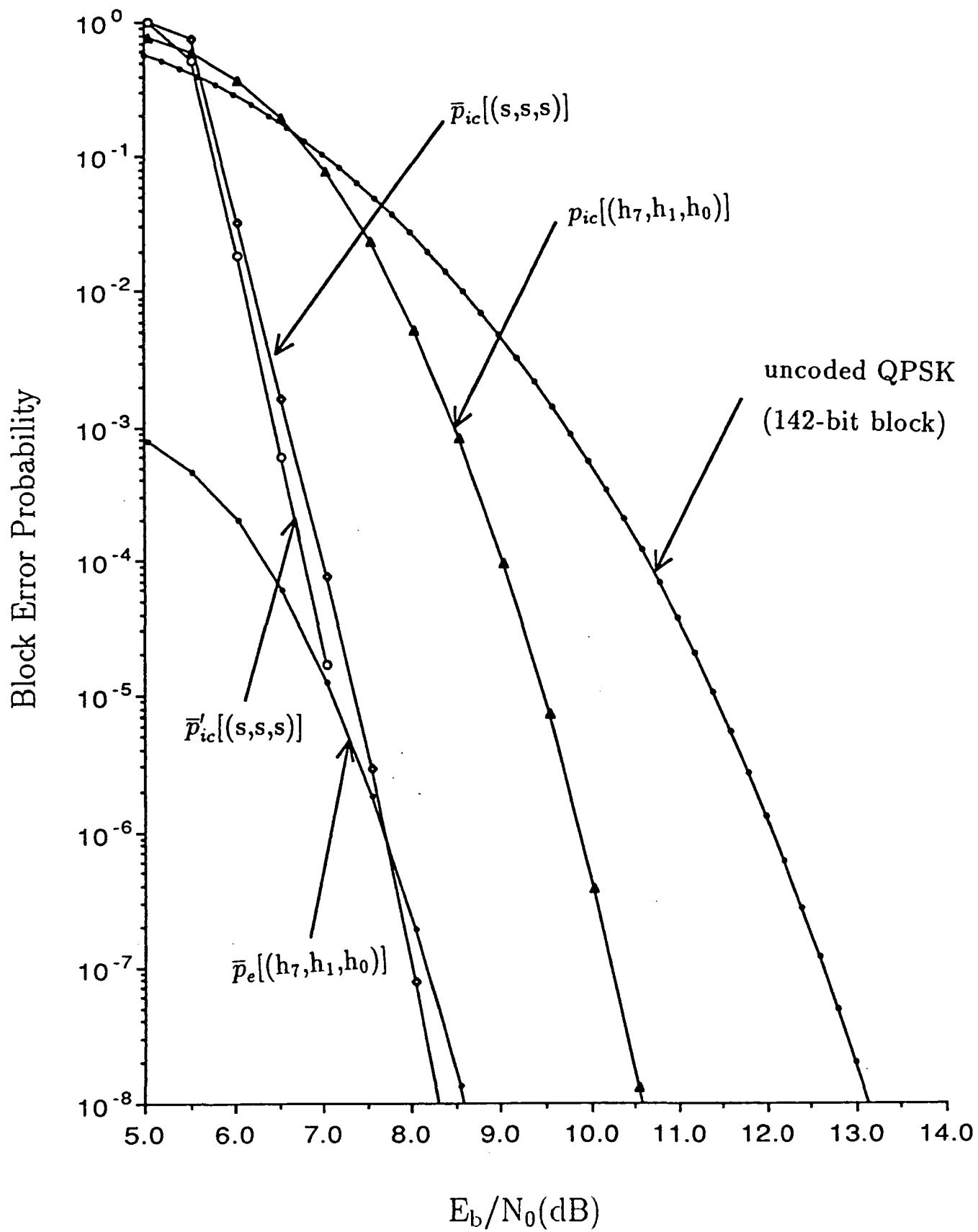


Fig.3. Error performance of the 8-PSK code  $RM_{6,2}RM_{6,4}P_{64}$ .

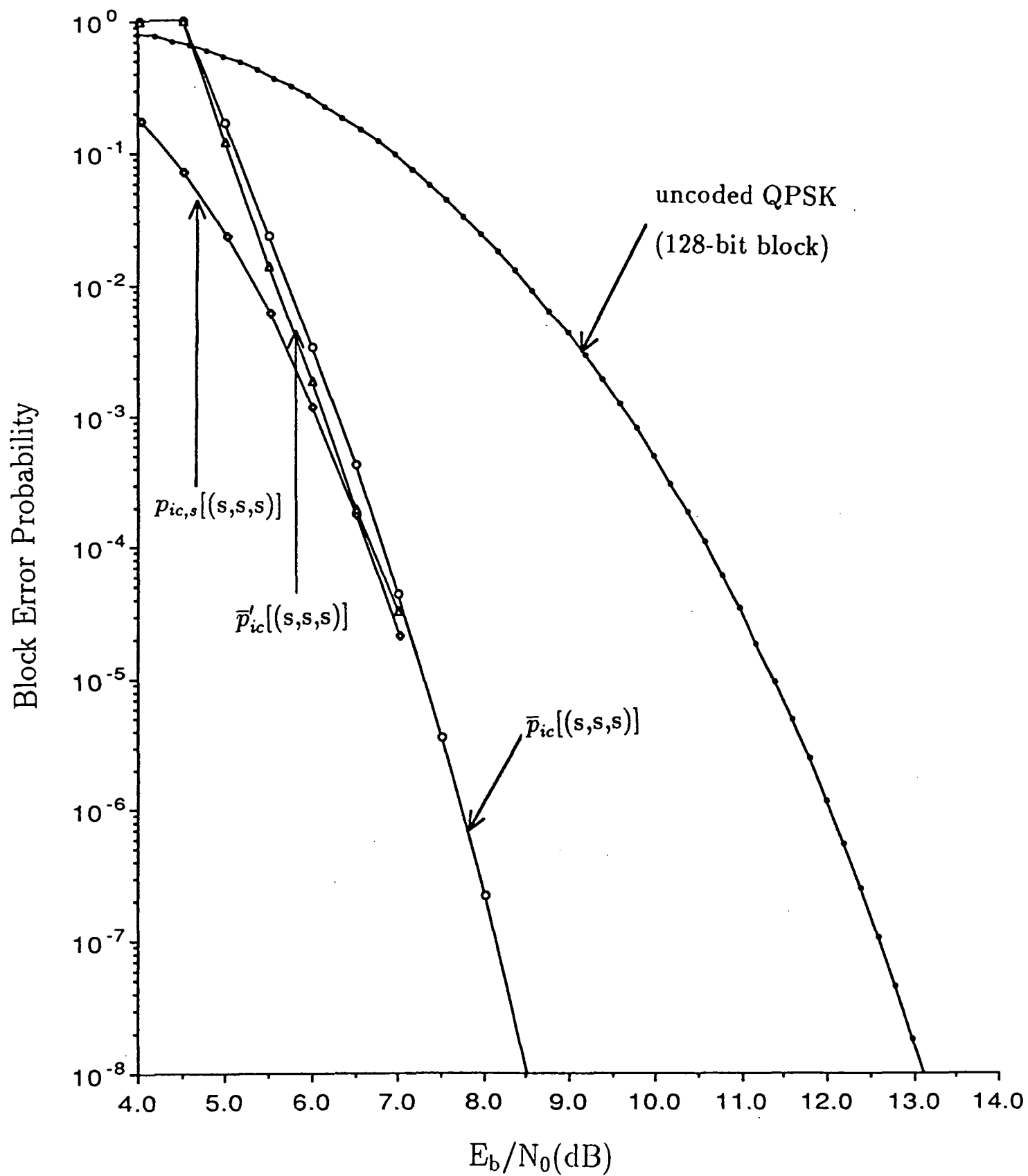


Fig.4. Error performance of the 8-PSK code  $RM_{6,1}RM_{6,4}P_{64}$ .

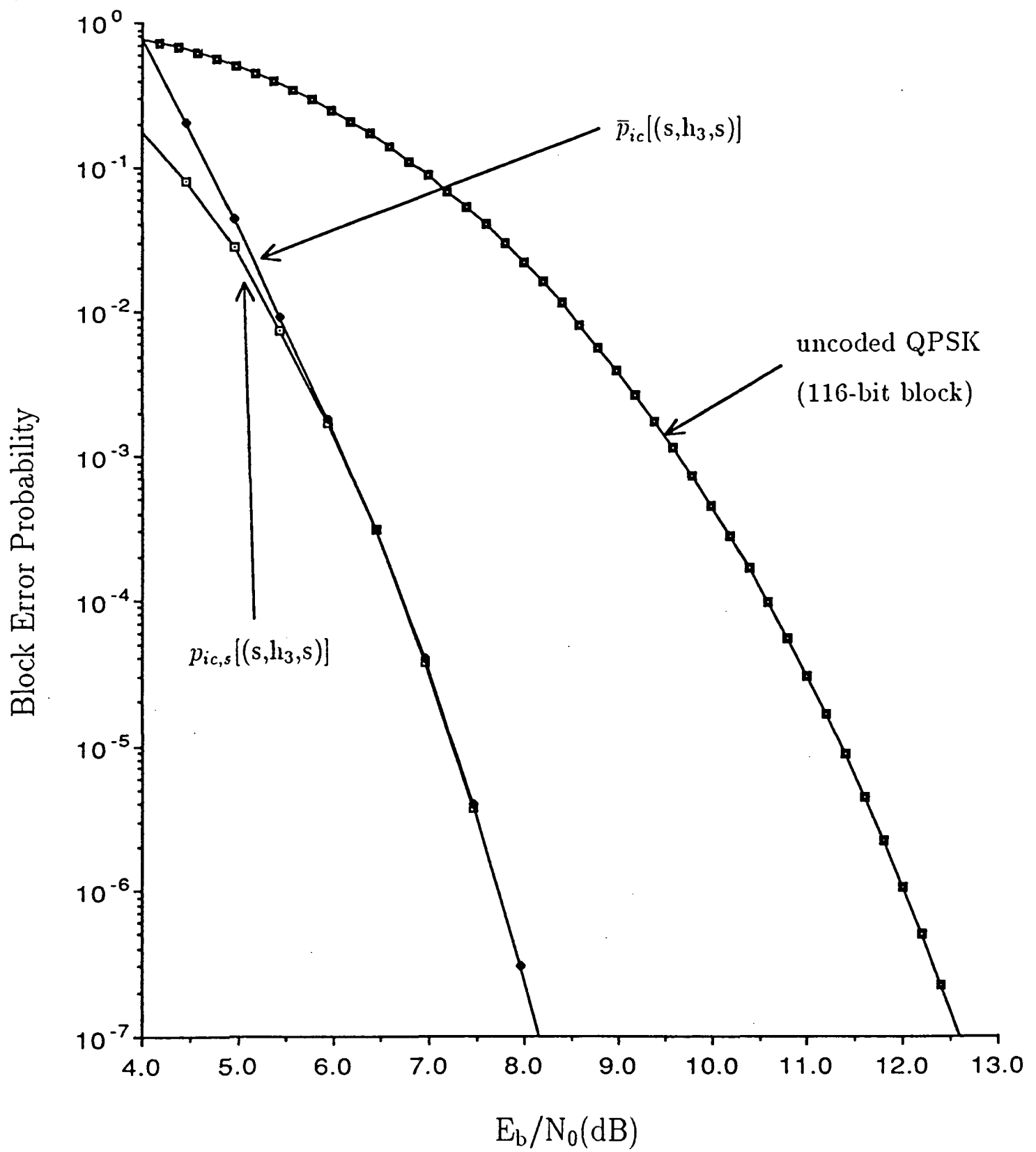


Fig.5. Error performance of the 8-PSK code  $RM_{6,1}ex-BCH_{6,7}P_{64}$ .

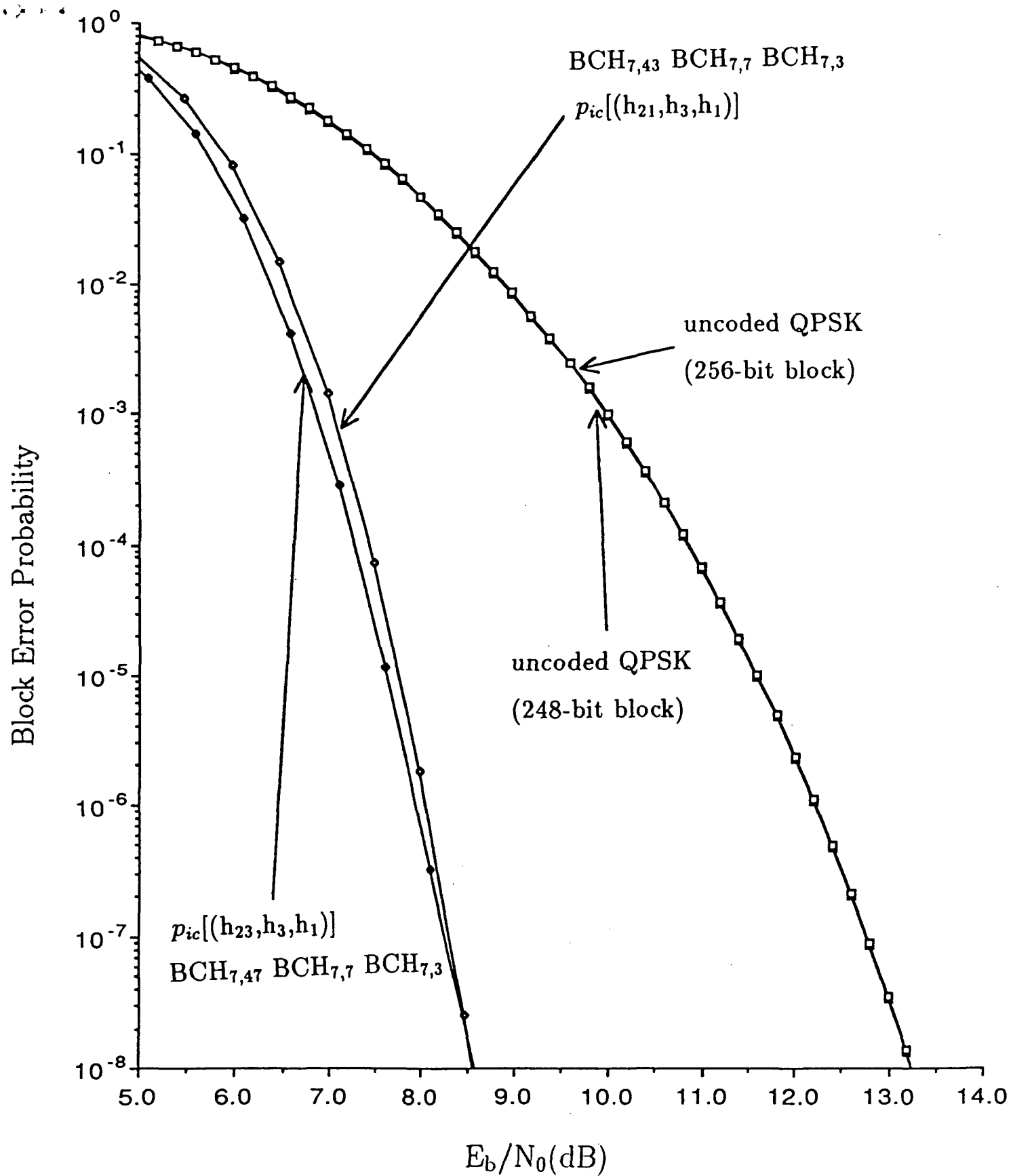


Fig.6. Error performance of the 8-PSK codes BCH<sub>7,47</sub>BCH<sub>7,7</sub>BCH<sub>7,3</sub> and BCH<sub>7,43</sub>BCH<sub>7,7</sub>BCH<sub>7,3</sub>

Table 1:

Some multi-level block modulation codes for which the error performances are evaluated for single-stage or multi-stage decoding

Notation	Signal Set	Definition of Codes	$n$	$R[C]$	$D_1, D_2, D_3$	Decoding Type	Complexity of Trellis Diagram
$C1$	8-PSK	$(RM_{5,1}RM_{5,3}P_{32})$	32	63/64	$4d_2$	$(3s)$	$2^9$
$C1$	8-PSK	$RM_{5,1}RM_{5,3}P_{32}$	32	63/64	$16d_1, 4d_2, 2d_3$	$(s, s, s)$	$2^4, 2^4, 2$
$C1$	8-PSK	$RM_{5,1}RM_{5,3}P_{32}$	32	63/64	$16d_1, 4d_2, 2d_3$	$(h_7, h_1, h_0)$	$-, -, -$
$C2$	8-PSK	$RM_{6,2}RM_{6,4}P_{64}$	64	142/128	$16d_1, 4d_2, 2d_3$	$(s, s, s)$	$2^{10}, 2^5, 2$
$C2$	8-PSK	$RM_{6,2}RM_{6,4}P_{64}$	64	142/128	$16d_1, 4d_2, 2d_3$	$(h_7, h_1, h_0)$	$-, -, -$
$C3$	8-PSK	$RM_{6,1}RM_{6,4}P_{64}$	64	127/128	$32d_1, 4d_2, 2d_3$	$(s, s, s)$	$2^5, 2^5, 2$
$C4$	8-PSK	$RM_{6,1}ex-BCH_{6,7}P_{64}$	64	115/128	$32d_1, 8d_2, 2d_3$	$(s, h_3, s)$	$2^5, -, 2$
$C5$	8-PSK	$BCH_{7,43}BCH_{7,7}BCH_{7,3}$	127	255/254	$43d_1, 7d_2, 3d_3$	$(h_{21}, h_3, h_1)$	$-, -, -$
$C6$	8-PSK	$BCH_{7,47}BCH_{7,7}BCH_{7,3}$	127	248/254	$47d_1, 7d_2, 3d_3$	$(h_{23}, h_3, h_1)$	$-, -, -$